# CLD3 Processes, Platforms & Examples

Aaron Schroeder, Research Associate Professor Social and Decision Analytics Division

> CLD3 Workshop March 25-26, 2019



**Biocomplexity Institute & Initiative** 

### CLD3 Processes, Platforms & Examples

- Introduce some of our platforms as they are used in our data science process
- Give some reasons why these platforms were selected
- Show some examples that highlight the use of these platforms at different stages of the process
  - Including an example of finding data and creating decision making tools for policy makers in Marshalltown
  - Scenario: Visualizations for Policy Makers to Assess Possible Transit Route Changes
  - Demonstrate
    - Ingestion of various administrative data sources using multiple method
      - job locations
      - vulnerable population locations
      - transit route locations
    - Creation of statistical indexes and maps for sub-county geographies





### VPS + Containerization = Data Ingestion Versatility!

- Simple Parallel
- Simple Additional IP Addresses (for cloud VPS services – use multiple NICs if running your own iron)
- Recent example: NSF project studying the value of Open Source Software
  - Needed many R installations to download just about every R and Python repository on Github!
  - Needed to determine the license being used on each using an online service licenses.io that limits the rate of use by your IP address. So needed multiple IP addresses





### Data Ingestion

- Establish type and method of data transfer
  - pushed to or pulled into the cooperative platform?
  - staying where it is and being dynamically queried in a federated manner as needed?
- Establish the best transfer protocol(s) to use given the types and method of transfer
  - e.g. SFTP, secure Dropbox, secured REST API, VT SAFR-Data Adapter for secure federated queries
  - Establish designed collection systems (e.g. behavioral experiments)
- Establish data marshaling processes
  - system mediation logic, data pipeline and data transformation, transfer schedule, and data provenance maintenance
- Establish secure data storage procedures
  - e.g. each project being stored on a new project-dedicated encrypted partition, original data being stored as nonremovable and non-editable

Establish Type of Access, Select Method, & Deploy Platform							
Туре	Ware	Federate					
Method	Push	Pull	Leave				
Protocol	SFTP, Secure Dropbox	Provider API, SAFR-Data Adapter					
Platform	Apach	NiFi, SAFR- Data Shaker					
Storage	Linux Ur Logical Vol	p (LUKS) nent (LVM)					



### Data Ingestion Libraries/Packages

#### library(tidytransit)

# get Marshalltown feed URL feed\_url <- feedlist\_df %>% setDT(.) %>% .[loc\_t %like% "Marshalltown, IA", url\_i]

# read gtfs data from url
gtfs <- read\_gtfs(feed\_url, geometry = TRUE, frequency =
TRUE)</pre>

🕏 gtfs	list [28] (S3: gtfs)	List of length 28
agency	list [1 x 7] (S3: spec_tbl_df, tb	A tibble with 1 rows and 7 columns
💿 areas	list [0 x 2] (S3: data.table, dat	A data.frame with 0 rows and 2 columns
calendar_attributes	list [4 x 2] (S3: data.table, dat	A data.frame with 4 rows and 2 columns
💿 calendar_dates	list [34 x 4] (S3: spec_tbl_df, 1	A tibble with 34 rows and 4 columns
💿 calendar	list [4 x 11] (S3: spec_tbl_df, 1	A tibble with 4 rows and 11 columns
O directions	list [0 x 3] (S3: data.table, dat	A data.frame with 0 rows and 3 columns
fare_attributes	list [1 x 7] (S3: spec_tbl_df, tb	A tibble with 1 rows and 7 columns
fare_rider_categories	list [0 x 3] (S3: data.table, dat	A data.frame with 0 rows and 3 columns
fare_rules	list [11 x 5] (S3: spec_tbl_df, 1	A tibble with 11 rows and 5 columns
farezone_attributes	list [0 x 2] (S3: data.table, dat	A data.frame with 0 rows and 2 columns
feed_info	list [1 x 10] (S3: spec_tbl_df, 1	A tibble with 1 rows and 10 columns
frequencies	list [0 x 5] (S3: spec_tbl_df, tb	A tibble with 0 rows and 5 columns
Iinked_datasets	list [0 x 7] (S3: data.table, dat	A data.frame with 0 rows and 7 columns
rider_categories	list [0 x 2] (S3: data.table, dat	A data.frame with 0 rows and 2 columns
routes	list [11 x 12] (S3: spec_tbl_df,	A tibble with 11 rows and 12 columns
💿 runcut	list [0 x 9] (S3: data.table, dat	A data.frame with 0 rows and 9 columns
Shapes	list [2174 x 5] (S3: spec_tbl_d	A tibble with 2174 rows and 5 columns
stop_attributes	list [13 x 2] (S3: data.table, da	A data.frame with 13 rows and 2 columns
stop_times	list [792 x 22] (S3: spec_tbl_d	A tibble with 792 rows and 22 columns
stops	list [76 x 15] (S3: spec_tbl_df,	A tibble with 76 rows and 15 columns
timetable_stop_order	list [0 x 5] (S3: data.table, dat	A data.frame with 0 rows and 5 columns
timetables	list [0 x 16] (S3: data.table, da	A data.frame with 0 rows and 16 columns
transfers	list [1 x 4] (S3: spec_tbl_df, tb	A tibble with 1 rows and 4 columns
trips	list [58 x 18] (S3: spec_tbl_df,	A tibble with 58 rows and 18 columns
stops_sf	list [76 x 14] (S3: sf, tbl_df, tk	A tibble with 76 rows and 14 columns
routes_sf	list [11 x 2] (S3: sf, tbl_df, tbl	A tibble with 11 rows and 2 columns
stops_frequency	list [169 x 6] (S3: data.table, (	A data.frame with 169 rows and 6 columns
routes_frequency	list [11 x 5] (S3: tbl_df, tbl, da	A tibble with 11 rows and 5 columns



### Data Ingestion APIs

	arlingtonva.us	Count	y Home • Terms of Use •	Policy • Feedback Datas	et Requests • Data Ca
PI / DEVELOPERS	HOUSING AND BUILI	DING			
Get your API key here.	<b>2019 Pro</b> Property and as	sessment information	(e.g. assessed amount	, tax balance)	
GUID: 2018-PROPE-ASSES-HISTO	IvwPropertyAssessmen	t ProvalLrsnld	RealEstatePropertyCode	MasterRealEstateProper	ReasPropertyStatusCod
Use it to invoke this Data View through the	74	134	01001007	01001007	Α
API. Learn more in the documentation.	111	136	01001009	01001009	Α
	148	137	01001010	01001010	A
API request examples:	185	139	01001012	01001012	A
Artrequest examples.	222	140	01001013	01001013	A
A normal invoke:	260	141	01001014	01001014	A
https://api.data.arlingtonva.us/ap i/v2/datastreams/2018-PROPE-ASSES- HISTO/data.json/?	334	143	01001016	01001016	A
	371	144	01001017	01001017	A
auth_key=YOUR_API_KEY&limit=50	408	145	01001018	01001018	Α
Change the output format:	445	147	01001020	01001020	Α
https://api.data.arlingtonus.ug/ap	482	148	01001021	01001021	Α
i/v2/datastreams/2018-PROPE-ASSES-	519	150	01001023	01001023	Α
HISTO/data.ajson/? auth key=YOUR API KEY&limit=50&	556	151	01001024	01001024	Α
	593	153	01001026	01001026	А
	5509	331	01006052	01006052	Α
Need help?	5547	332	01006053	01006053	А
Send us an e-mail. We are always happy to	5584	333	01006054	01006054	Α
help!	5621	334	01006055	01006055	А
	5658	336	01006057	01006057	Α
	5695	337	01006058	01006058	A
	5733	339	01006060	01006060	Α

## </>

i

 $\bigcirc$ 

 $\forall$ 

 $\uparrow_{\downarrow}$ 

\_\_\_\_7

### Ingest Data: Collection, Experiment (oTree)

	Concerning Helps (1/12/2000)	essions and fo stan valet espine et carja	Espe or se d tri tabel	Pages compared 1/10 pages	Ionohe Comohe Comentae parte parte	do (der b Norder 1 1 1 1 1 1 1 1 1 1 1 1 1	mo) Presults Corrent page name Presolution Corrent of Corrent of Corren	S Po Status Paying Paying	o L	C Relation			
<section-header>         Descented on sension include (sense)         Image: Imag</section-header>	Dashbor Comp Stinsession deplay P1 P2 Address stores Introduction Introduction Stinsess	Code visit sagende	Ø Crit	Pages completed 1/10 pages	Current app name	do (der landor Round Round 1 1 1 1 1 1 1 1 1 1 1 1 1	mo) Create Anna Peroduction Control (	S Po Status Playing Playing	ayments Last request succeeded True True	C Relat			
Image: Imag	Line section display P1 P2 Advance sizes Introduction Instructions Via have been rendering and anny will be selected at renderin to be pair will be selected at renderin to be pair will be selected at rendering to be series	Code viciteit sagarute	8 Lin	Pages completed 1/10 pages	For Current app name	Monter     Round     number     1     1	E Pesuta Current page Asima Introduction	\$ P	ayments Lest request succeeded True	C Receil			
Intermediation in the state is the state	Lit in session display P1 P2 Advance sidner Introduction Instructions Yas have been rendomly and array will be asked at random to be pair will be asked at random to be pair will be asked at random to be pair and the asked at random to be pair asked at random to be pair asked at random to be pair.	Code Valet sogerule	Label	Pages comparised 1/10 pages	Current apo name	Round number 1 1	Current page name Introduction Constant	Statue Paying Playing	Last request auccended If us The	Levi page timestamp less than a minute ago			
principal pr	Introduction Instructions You have been strateging and arrays with be selected at modern to be par with be selected at modern to be participant A	sagarule at userjaj		010 Jages	P	1 1	Comment Comment	Paying	1.4	initia ago			
<text><section-header><section-header><section-header><text><section-header><text><text><text><text><text></text></text></text></text></text></section-header></text></section-header></section-header></section-header></text>	Abance stowe Introduction Instructions You have been rendomly and array will be elected at moden to be par- will be elected at moden to be par-	at usarijaj				ur							
	Introduction Instructions You have been randomly and aroung will be seeked at random to be participant A				P	ur	•						
<section-header><section-header><section-header><text><text><text><text><text></text></text></text></text></text></section-header></section-header></section-header>	Instructions Instructions You have been rendomly and arony will be askeded at rodom to be per learn whether you are participant A					Ur							
	To star, participant A monotwen 100, can send some cal of the 100 point they will be tripled. Once it receives of this point to A. For your convenience, these trating screens of this study. On the next page, you will have to an instructions correctly.	Infocurations  Infocurations  Infocurations  Infocurations  Infocuration  Infocuration						Understanding Question 1 of 1  Suppose that participant A sent 20 points to participant B. Hown served the tops amount, participant B sent 30 points to participant A. In the end, how many points a participant A would have:  points  Purticipant B would have:  points  Mane  Instructions  You have been randomly and anonymously pained with another participant. One of all to selected of random to be participant A, the other will be participant B. You same whether you as participant A. You chanse tools not analog and decision.  To start, participant A another 100 points to participant B. Mouse provide the result of the participant A. The other will be participant B. You same whether you as participant A. In the participant B. You same whether you as participant B. Y					

• oTree lets you create:

- Controlled behavioral experiments in economics, market research, psychology, and related fields
- Multiplayer strategy games, like the <u>prisoner's dilemma</u>, <u>public goods</u> <u>game</u>, and <u>auctions</u>.
- <u>Surveys</u> and quizzes, especially those that require customized or dynamic functionality not available with conventional survey software.



## Data Information Process & Platform (Lexicon / PostgreSQL)

• The Lexicon: an inventory of and history of changes to:

- -every available data field in every available data source
- -the structure of their storage
- -possible values and meanings of the information
- -possible transformations of each set of field values from one data source to another another data source
- -methods of data source access
- -matching algorithms and how they are to be used in conjunction with possible field value transformations
- Provides fundamental functions for the operation of the framework and is a requirement that the data information be collected from all partner communities
- Enables removal of much complexity required for high quality data linkage
  - i.e. No enforcing data standardization schemes on data partners



### Data Information Process & Platform

#### Lexicon ER Diagram

	metadata	_table
Key	table_name	
	friendly_name	
	description	
	critical_changes	
	last_updated	

	metadata_valid_valu	les
Key	table_name	
Key	column_name	
	value	
	description	
	valid_use_begin_date	
	valid_use_end_date	
	last_updated	

	metadata_trar	nsformation
Key	table_name	
Key	column_name	
	description	
	transformation	SQL code
	ordinal	
	target	value   metadata
	last_updated	

	metao	data_column
Key	table_name	
Key	column_name	
	friendly_name	
	description	
	requirements	required   conditional   not required   required before/after date
	demographic_type	race/ethnicity   gender   dob/age   SES   cohort
	domain_type	categorical   numeric   date   ID   open
	numeric_range_min	
	numeric_range_max	
	date_range_min	
	date_range_max	
	id_length_min	
	id_length_max	
	critical_changes	changes in domain (e.g. valid values, evidenced by appearance of new values); changes in collection method (e.g. change to forced-selection from free text, evidenced by frequency of non-entry and error rates)
	valid_use_begin_date	
	valid_use_end_date	
	original_collection_source	where collected originally and why
	original_entry_by	subject   operator   auto-generated
	last_updated	

### Data Storage and Management

Database Choice: PostgreSQL / PostGIS Database Interfaces: Adminer, Rstudio, Jupyter, psql

## GIN Indexes + trigrams! ©

Schule Case Case Case Case Case Case Case Cas	Ctructur	cod &	llnctr	ucturo	d Dati	~		Language: English	PostgreSQL » postgis » sdad » geospatial\$census_tl » Select: tl_2018_19_block_centerpoints
Image:       Point rest:	Structur	eux	UIISU	ucture		ג		Adminer 4.6.3 4.7.1	Select: tl_2018_19_block_centerpoints
Additional 4.3.4.3.4.3.4.3.4.3.4.3.4.3.4.3.4.4.4.4	Language: English	PostgreSQL » postg	s » sdad » geospatial\$cen	sus_cb » Select: cb_2016_01_	bg_500k			DB: sdad	Select data Show structure Alter table New item
Bit will will will will will will will wi	Adminer 4.6.3 4.7.1	Select: cb_2	016_01_bg_500k					geospatial\$census_ti	Select Search Sort Limit Select Select
Start:       Star:       Start:       Start:	DB: sdad	Select data Sho	w structure Alter table	New item				SQL command Import Export Create table	SELECT * FROM "t1_2018_19_bink_centerpoints" LINIT 50 (0.038 s) Edit
Image: Control         Image: Contro         Image: Control         Image: C	Schema:	Colort Corre	Cant Lincit	Tout log other Antion				select geography_columns	Modify geoid geometry
Bit       B	geospatiai\$census_cb			Text length Action				select geometry_columns select raster_columns	edit 1901300300 035 0101000020E6100000DF2B0483B51D57C000304388D02D4540
Ender         Cardinal Construction         Cardinal Co	SOL command Import		50	100 Select				select raster_overviews	
estect 0: 2016. 01. 02. 50/k               method             meth	Export Create table	SPIRCE + FROM Sch 20	16 01 bg 500k" TIMTE 50 (0.04	1 o Edit				select spatial_ref_sys	edit 1901300290 090 0101000020E6100000EF5DDE77E70E57C05401F73C7F2E4540
select         2016         0.1.bg. 2004         0.0000         1.000000000000000000000000000000000000		billion CD_20	10_01_09_000x DIMI1 00 (0.04					select tl 2017_us_county	edit 190130028000 97 01010000004FA9B7AB7C0857C0C1E09A3BFA2D4540
select         0.00         <	select cb_2016_01_bg_500k	Modify STATEFF	COUNTYFP TRACTCE	BLKGRPCE AFFGEOID	GEOID NA	ME LSAD	ALAND	select tl_2018_19_bg_centerpoints	edit 1901300280030 10101000020E610000017EBC0DE690657C0012E6DDD722F454
select         0.10         <	select cb_2016_01_county_within_u	<b>edit</b> 01	077 011501 5	5 150000US0107701	15015 010770115015 5	BG	6844991	select tl_2018_19_block_centerpoi	edit 190130028003095 11000020E510000018D75306690657C0C754807C4C2F
select         2016         0.1         0.0	select cb_2016_01_cousub_500k	<b>edit</b> 01	045 021102 4	4 150000US0104502	11024 010450211024 4	BG	1136085	select tl 2018 51 bg	edit 190130028033094 010. 920E61000002C5382B4B50657C01568C1DF 474540
select         2.016_01_pure10_500c         eff         0         0         0         0.00000000000000000000000000000000000	select cb_2016_01_place_500k	🗆 edit 01	055 001300 3	3 150000US0105500	13003 010550013003 3	BG	1378742	select tl_2018_51_bg_block_center	edit 190130028003093 0101000020. 5000114C0035905572 aD900274540
select         2.01.2         0.1         0.99         0.400         1         1.000000000000000000000000000000000000	select cb_2016_01_puma10_500k	<b>edit</b> 01	089 001700 2	2 150000US0108900	17002 010890017002 2	BG	1040641	select tl_2018_51_bg_centerpoints	edit 19013002803104 0101000020E61000004B619456B20557C013245B02952E4540
select         2016_01_rrst_500K         edt         01         0000         4         15000000016 (0000001000010000100000100000100000100000	select cb_2016_01_sldl_500k	o edit 01	069 041400 1	1 150000US0106904	14001 010690414001 1	BG	8243574	select tl_2018_51_tabblock10	edit 190130028003113 0101000020E61000000AC09A5E300457C090BA42C4722D4540
select         0.016         0.01         0.05         3         150000005010000202         0.101000022         0.1010000251000000510000005100000005100000000	select cb_2016_01_sidd_500k	odit 01	073 010801 4	4 150000US01073010	08014 010730108014 4	BG	1303598		edit 190130020002016 0101000020E61000004FEB9161701057C0C375D6B8483B4540
select 0, 2016, 0.2, arr. 500k       edt       01       015       00200       1       19000000010000000000000000000000000000	select cb_2016_01_unsd_500k	edit 01	101 005102 3	3 150000US0110100	51023 011010051023 3	BG	677515		edit 190410801001145 0101000020E510000081D71041E0C05/C0B8+8252C3B794540
select - 2016 - 02 - 0000-00000000000000000000000000	select cb_2016_02_anrc_500k	edit 01	015 000200 1	150000US0101500	02001 010150002001 1	BG	4085127		edit 190410801001144 0101000020E6100000457F68E6C9C057C08C648F5033794540
select 0.2016 02 county 500k         edt         01         095         03100         3         1500000050931003         0.995031103         3         66         126376           select 0.2016 02 place 500k         edt         01         097         00301         2         15000000501000015030         1000000261000000785035301147C0082530333541457C00845769353035440           select 0.2016 02 place 500k         edt         01         037         01200005110010         101000002610000078530335511437C0084540           select 0.2016 02 place 500k         edt         01         033         15000000051000101010101002         2         66         85839           select 0.2016 04 county within 1         edt         01         01400         2         1500000051000100101010002         101000002610000074502000150001050750070750025400           select 0.2016 04 county within 1         edt         01         01400002         2         15000000510000050100000000000000000000	select cb_2016_02_bg_500k	edit 01	069 041900 2	2 150000US0106904	19002 010690419002 2	BG	2032290		edit 190410804002009 0101000020E51000004CEC905104CB5/C08AA9995A5B9F4540
select cb         Cl         Co         Co        Co        Co <t< td=""><td>select cb_2016_02_cousub_500k</td><td>edit 01</td><td>095 031100 3</td><td>3 1500000US0109503</td><td>11003 010950311003 3</td><td>BG</td><td>1263378</td><td></td><td>edit 190410801001155 0101000020E610000061BB20C778C157C09E3325D4A764540</td></t<>	select cb_2016_02_cousub_500k	edit 01	095 031100 3	3 1500000US0109503	11003 010950311003 3	BG	1263378		edit 190410801001155 0101000020E610000061BB20C778C157C09E3325D4A764540
select C         0.116_0.7_glunal10_SUDR         edit         01         073         012305         4         100300123052         0         1003002061200000000000000000000000000000	select cb_2016_02_place_500k	edit 01	097 003901 3	2 15000000000000000000000000000000000000	39012 010970039012 2	BG	1183703		edit 190130026011076 0101000000F95F338C181A57C0A848CF3EA03C4540
select ch. 2016 0.2 selful \$00k       0.1       0.3       0.12003       4       15000000501030115013       110000021011       1010000220E510000027012802(2150000750114540)         select ch. 2016 0.2 selful \$00k       edit       01       025       011501       2       15000000501030115013       15000000501030115013       1500000050102011010100220E510000027E510000027E510000027E510000027E510000027E510000027E510000027E510000027E510000027E510000027E510000027E510000027E510000027E510000027E510000027E510000027E51000007E5132017C07408028344540         select ch. 2016 0.4 selful \$01       01       097       007600       1       15000000501070010011       10770070001       1       BG       1316209         select ch. 2016 0.4 selful \$00k       edit       01       097       007600       1       150000005010700700001       1       BG       1316209         select ch. 2016 0.4 selful \$00k       edit       01       097       007600       1       1500000050107000700502       BG       550870         select ch. 2016 0.4 selful \$00k       edit       01       073       010001       1       150000005010700300000       186       550870         select ch. 2016 0.4 selful \$00k       edit       01       073       003600       1       150000005010900073030000       3       BG       500888         select ch. 2016 0.4 se	select cb_2016_02_puma10_500k	edit 01	072 012205	150000000000000000000000000000000000000	23054 010730123054 4	BC	2252791		edit 190130015022052 0101000020E61000038E04E7DFB1957C0368BBC51D03C4540
select cb       2016       02       v1       003       011301       3       100000005011201       005       011301       2       150000005011201       012       6       6858339         select cb       2016       04       01       101       001400       2       1500000050112011001       2       8       6858339       edit       101       0014000266100000748494892175/70074A008469444540         select cb       2016       04       01       011       001400       2       15000000501100101002       2       86       899418         select cb       2016       04       010       077       07600       1       15000000501070070001       1097076000       1       86       1010002065100000757257C7105700175802874745544544         select cb       2016       04       01       073       014000       2       1073000000       1073000000       1073000000       1073000000       1073000000       1073000000       1073000000       1073000000       1010002065100000578170705700570057004808274645450         select cb       2016       04       01       073       03600       1       150000000501073000000       1       86       50883       edit       190100002065100000058187705707057020469834654540       edit	select cb_2016_02_sldu_500k	edit 01	002 011501 3		15012 010020115012 2	BG BC	2332701		edit 199130020002011 010100000261000004701A260C61057C0B07E029A303C4540
select ch_2016_02_unsd_500k       0.01       1.1       1.1       1.1       0.019001       2       150000005011150119012       2       86       858339         select ch_2016_04_county_within_1       edit       0.1       0.01400       2       1500000050101001002       0.011001002       101000020E510000000F32138912070756886254143490         select ch_2016_04_county_within_1       edit       0.1       0.073       0.07600       1       15000000501070010002       1       86       316209         select ch_2016_04_elsd_500k       edit       0.1       0.73       0.04000       2       15000000501070010002       2       86       39870         select ch_2016_04_elsd_500k       edit       0.1       0.73       0.04000       2       1500000050107001002       2       86       39870         select ch_2016_04_elsd_500k       edit       0.1       0.73       0.04000       2       150000005010700002       2       86       39870         select ch_2016_04_elsd_500k       edit       0.01300020261000005761877017507080228464540       0.01070020661000005761877015700582745454540       0.0100002061000005761877016700570827245464540       0.010000206100000576187711570054928245545464540         select ch_2016_04_elsd_500k       edit       0.01       0.030       3       15000000	select cb_2016_02_tract_500k		125 011001		10010 011000110010 0	DG DC	2322/43		edit 190130004001011 0101000020E610000024BF34FB22157C071F5AD2017414540
select CD       01       01       01       001 <t< td=""><td>select cb_2016_02_unsd_500k</td><td></td><td>125 011901 2</td><td></td><td>19012 011250119012 2</td><td>BG</td><td>000440</td><td></td><td>edit 190130019001001 10101000020E5100000DF412A2D7125/C0E7DE686C54414540</td></t<>	select cb_2016_02_unsd_500k		125 011901 2		19012 011250119012 2	BG	000440		edit 190130019001001 10101000020E5100000DF412A2D7125/C0E7DE686C54414540
select b. 2016 0.4_cousub_SOOK       edit       01       073       01001       1       15000000S0109/00/0001       101900020E10000027850284F105/C01058962F364540         select b. 2016 0.4_elds 500k       edit       01       073       01000       2       15000000S0109010101       1073       01000020E510000027850284F105/C01058962F364540         select b. 2016 0.4_elds 500k       edit       01       073       01400       2       15000000S0109000002       2       86       344082         select b. 2016 0.4_elds 500k       edit       01       097       003602       2       15000000S01090030002       2       86       55822         select b. 2016 0.4_elds 500k       edit       01       097       003600       3       15000000S01090300003       1089030003       3       86       500888         select b. 2016 0.4_elds 500k       edit       01       039       962500       1       1500000US010399052001       1       86       127235         select b. 2016 0.5_ext 500k       edit       01       039       962500       1       1500000US010399052001       1       86       127235         select b. 2016 0.5_ext 500k       edit       01       039       962500       1       1500000US0101000000000000000000000000000	select cb_2016_04_bg_500k	edit 01	101 001400 2	2 150000050110100		BG	899418		edit 190130026035002 01010000015AC71369D1F57C047A00B34E9444540
select cb. 2016, 04_elsd_s00k       edit       01       073       01000       1       1500000US010730110011       1       BG       350870         select cb. 2016, 04_elsd_s00k       edit       01       073       00400       2       1500000US01073014000       20173014002       2       BG       34482         select cb. 2016, 04_elsd_s00k       edit       01       073       003602       2       1500000US01073004002       2       BG       34482         select cb. 2016, 04_eldid, 500k       edit       01       089       03000       3       BG       500888         select cb. 2016, 04_eldid, 500k       edit       01       073       03600       1       150000US0107300600       1073036001       1       BG       371522         select cb. 2016, 04_eldid, 500k       edit       01       039       962500       1       150000US01039625001       10399625001       1       BG       617235         select cb. 2016, 05_eldis, 500k       edit       01       01       019       020201025       1010000261000002471880457045430         select cb. 2016, 05_eldis, 500k       edit       01       01       01       010900055       S       S       S       S         select cb. 2016, 05_bldiso	select cb_2016_04_cousub_500k		097 007600 1	1 1500000050109700	/6001 0109/00/6001 1	BG	1316209		edit 190130022001024 0101000020E5100000/A56D28A6F1D5/C01/36B9CE/3594540
select Cb 2016 04 place 500k select Cb 2016 04 puma10 500k select Cb 2016 04 scsd 500k       edit       01       073       004000 2       2       1500000US01730040002 12       BG       344682         select Cb 2016 04 puma10 500k select Cb 2016 04 scsd 500k       edit       01       097       00302 2       1500000US010730040002 12       BG       665822         select Cb 2016 04 scsd 500k       edit       01       073       00300       1       1500000US01073004000 10073003002       2       BG       659822         select Cb 2016 04 scsd 500k       edit       01       073       00300       1       10030020610000576CCB081557C0BF2ABI22794540         select Cb 2016 04 urat 500k       edit       01       073       03600       1       1003099625001       1       BG       61127235         select Cb 2016 05 bg 500k       edit       01       01890       02721       2       1500000US0173009005       1       BG       61127235         select Cb 2016 05 countywithin 1       01       01890       02721       2       1500000US01730109005       1       BG       2162239         select cb 2016 05 countywithin 1       01       073       01400       1       010301020212       0100002061000000231618747005557706552700555770655770655770655770655770655770655770655770655770655	select cb_2016_04_elsd_500k	edit 01	073 011001	1 1500000050107301	10011 010730110011 1	BG	350870		edit 190130022001025 0101000020E1000000EC/F25FC//105/C010595AE8/545450
select cb_2016_04_scsd_500k       edit       01       097       003002       2       1500000US01097003602       101097003602       2       86       695822         select cb_2016_04_scd_500k       edit       01       089       00300       3       1500000US01097003602       1089003003       3       86       50088         select cb_2016_04_sidl_500k       edit       01       039       962500       1       1500000US0109003003       1089030003       3       86       50088         select cb_2016_04_sidl_500k       edit       01       039       962500       1       1500000US010399625001       10730036001       1       86       6112235         select cb_2016_04_sidl_500k       edit       01       039       962500       1       1500000US01010018001       1       86       6112235         select cb_2016_05_sousb_500k       edit       01       01       0180       01       10100002212       10030002121       01399625001       1       86       621192         select cb_2016_05_sousb_500k       edit       01       01       01       01       073       100000       5       103000002712       2       86       511192         select cb_2016_05_sousb_500k       edit <t< td=""><td>select cb_2016_04_place_500k</td><td>edit 01</td><td>073 004000 2</td><td>2 150000US01073004</td><td>40002 010730040002 2</td><td>BG</td><td>344082</td><td></td><td>euit 190130026042055 0101000020E5100000E32215B12D175/C0F4BD282746464540</td></t<>	select cb_2016_04_place_500k	edit 01	073 004000 2	2 150000US01073004	40002 010730040002 2	BG	344082		euit 190130026042055 0101000020E5100000E32215B12D175/C0F4BD282746464540
select b. 2016 04_sidl_500k       edit       01       089       03000       3       1500000US010890030003       10       BG       500888         select b. 2016 04_sidl_500k       edit       01       073       03600       1       1500000US010890030003       1073003000       1073003000       1       BG       500888         select b. 2016 04_sidl_500k       edit       01       073       03600       1       1500000US01039962500       1       BG       571522         select b. 2016 05_05,500k       edit       01       019       020202       1       1500000US01039962500       1       BG       6117235         select b. 2016 05_05,500k       edit       01       089       002721       2       10080002712       2       BG       21289         select b. 2016 05_county_within_1       select b. 2016 05_county_within_1       select b. 2016 05_county_within_1       1500000US01089002712       2       BG       21289         select b. 2016 05_place_500k       edit       01       073       01900       1       1000000S01089002712       2       BG       21289         select b. 2016 05_place_500k       edit       01       073       01300       1       BG       101909020E100000010000000000000000000000	select cb_2016_04_scsd_500k	edit 01	097 003602 2	2 150000US01097003	36022 010970036022 2	BG	695822		edit 190130017022005 01010000055CCCEB081557C086C2A3E8DC434540
select Cb 2016_04_sidu_500k       edit       01       073       003600       1       1500000050173003000       10730036001       1       BG       371522         select Cb 2016_04_unsd_500k       edit       01       039       962500       1       1500000050198/07/03962500       1       BG       617235         select Cb 2016_05_unsd_500k       edit       01       010       0101       0101000110100000       10100000261000000348701057C0546474540         select Cb 2016_05_county_within_       selet Cb 2016_05_place_500k       edit       01       073       01900       5       1000000501073003000       1       BG       21129         select Cb 2016_05_county_within_       selet Cb 2016_05_place_500k       edit       01       073       01900       5       1000000501073003000       1       BG       21289         select Cb 2016_05_place_500k       edit       01       073       01900       5       1000000501073001000       1       BG       12189         select cb 2016_05_place_500k       edit       01       073       01300       1       BG       10199       5       BG       12189         select cb 2016_05_place_500k       edit       01       073       01300       1       BG       1069800	select cb_2016_04_sldl_500k	<b>edit</b> 01	089 003000 3	3 150000US01089003	30003 010890030003 3	BG	500888		eur 19041081002188 1101000020e1000005/88D60885C5/C00BAPDB122/94540
select cb 2016 04 unsd 500k       edit       01       039       962500       1       15000000510399625001       1       B6       612723         select cb 2016 05 bg 500k       edit       01       010       010002       1       1500000051010000800       10100002       101000024651000008087/100700647/94540         select cb 2016 05 county_within _ 073       01000000510730109005       10089002721       2       B6       2126239         select cb 2016 05 county_within _ select cb 2016 05 county_within _ select cb 2016 05 county_within _ 073       15000000510730109005       10730109005       5       86       512189         select cb 2016 05 county_within _ 01       01       073       01300       1       0100000000000000000000000000000000000	select cb_2016_04_sidu_500k	<b>edit</b> 01	073 003600 1	150000US0107300	36001 010730036001 1	BG	371522		
select cb. 2016.05_bg_500k       edit       01       101       001800       1       15000000501101010800       10110108002       1101080002       11011150       1101080002       11011150       11011080002       11011150       11010108000000000000000000000000000000	select cb_2016_04_unsd_500k	<b>edit</b> 01	039 962500 1	150000US0103996	25001 010399625001 1	BG	6127235		edit 199410801001113 010100020E1000003AD8710070C157C0F04044540
select cb_2016_05_county_within 1       edit       01       089       002721       2       1500000010108000027212       010890027212       2       BG       212623         select cb_2016_05_place_500k       edit       01       073       010900       5       15000000010730109005       5       BG       512189         edit       01       073       013400       1       1500000001730109005       5       BG       512189         edit       01       073       013400       1       150000000173013400       101730134001       1       BG       106989         select cb_2016_05_sidl_500k       edit       01       077       010700       2       1500000001770107002       2       BG       749731       0       010600002001000000010000000000000000000	select cb_2016_05_bg_500k	<b>edit</b> 01	101 001800 1	150000US0110100	18001 011010018001 1	BG	2611192		
Select Cb 2016 05_ploace_500k       edit       01       073       010900       5       1500000010730109000       5       BG       512189       12 3 4 5 last       Induity       Select Cb 216,05 ploace       Select Cb 2016 05 ploace       Select 2016 05 ploace       Select 201	select cb_2016_05_county_within_u	<b>edit</b> 01	089 002721 2	2 150000US01089002	27212 010890027212 2	BG	2126239		Page =
select cb 2016_05_puma10_500k select cb 2016_05_sidu_500k select cb 2016_05_sidu_500k	select cb_2016_05_cousub_500k	🗆 edit 01	073 010900 5	5 150000US0107301	09005 010730109005 5	BG	512189		rage mole result mole result (* 210,007)
select cb_2016_05_sldL_500k select cb_2016_05_sldu_500k	select cb_2016_05_puma10_500k	<b>edit</b> 01	073 013400 1	150000US0107301	34001 010730134001 1	BG	1006985		1 2 3 4 5 last A 210,007 rows Save Leat Clone Delete
	select cb_2016_05_sldl_500k	edit 01	077 010700 2	2 150000US0107701	07002 010770107002 2	BG	749731	0 0106000020AD100000010	0000001030000001000000000000015527E52EDE9550
select cb 2016 05 tract 500k	select cb_2016_05_sidu_500k	edit 01	101 002700 1	1 150000US0110100	27001 011010027001 1	BG	1278063	0 0106000020AD100000010	000000103000000010000000000000000000000
select cb_2016_05_unsd_500k Page Whole result Modify Selected (0) Export (3,437)	select cb_2016_05_unsd_500k	Page	Whole result	fy Selected (0)	-Export (3,437)				
select cb. 2016_06_bg_500k 12.3.4.5 69 3,437 rows Save Edit Clone Delete	select cb_2016_06_bg_500k	1 2 3 4 5 69	3,437 rows Save	Edit Clone Delete					

### Store most used geographic places

Language: English	PostgreSQL >	» postgis »	sdad » geo	spatial\$places » Select: us_pl_urgent_care_facilities				Logout
Adminer 4.6.3 4.7.1	Select:	us_pl_u	urgent_	care_facilities				
DB: sdad 📀 Schema:	Select data	show st	tructure	Alter table New item				
geospatial\$places	Select	-Search-	Sort	Limit Text length Action				
SQL command Import Export Create table	SELECT * FROM	M "us_pl_urg	jent_care_fa	cilities" LIMIT 50 (0.046 s) Edit				
select us_pl_cellular_towers	Modify C	<b>DBJECTID</b>	ID	NAME	TELEPHONE	ADDRESS	ADDRESS2	СІТҮ
select us_pl_colleges_universities	edit 4	001 1	11513140	ARRAGUT WALK-IN CLINIC	865-671-6026	11408 KINGSTON PIKE		KNOXVILLE
select us_pl_day_care_centers	edit 4	002 1	10422042	RISTATE URGENT CARE OF OAKLEY	513-531-1505	5002 RIDGE AVENUE		CINCINNATI
select us_pl_local_law_enforcement	edit 4	003 1	10425970	OLANTIC WALK-IN URGENT CARE - MANDARIN	904-288-0277	12303 SAN JOSE BOULEVARD		JACKSONVILLE
select us_pl_mobile_home_parks	edit 4	004 1	10844061 (	ONCENTRA URGENT CARE - HARRISBURG WEST	717-795-1819	4910 RITTER ROAD		MECHANICSBURG
select us_pl_nursing_homes	edit 4	005 1	11513174	L CENTRO HERIDAS Y ULCERAS	787-735-8001	CALLE JOSE C VAZQUEZ		AIBONITO
select us_pl_places_of_worship	edit 4	006 1	10844294	OLANTIC WALK-IN URGENT CARE - POMPANO BEACH	954-580-4401	1611 SOUTH FEDERAL HIGHWAY		POMPANO BEACH
select us_pl_private_schools	edit 4	007 1	10469731 (	ARDENS URGENT CARE	561-626-4878	3555 NORTHLAKE BOULEVARD		WEST PALM BEAC
select us_pl_public_schools	edit 4	008	10470864	IEDEXPRESS URGENT CARE - SOUTH HILLS	412-854-3627	2600 OLD WASHINGTON ROAD		UPPER SAINT CLA
select us_pl_snap_providers select us_pl_urgent_care_facilities	edit 4	i009 1	11513712	OUSTON MEDICAL CENTER - PAVILLION MED STOP	478-923-2843	233 NORTH HOUSTON ROAD		WARNER ROBINS
select us_pl_va_medical_facilities	edit 4	010	11513134	JOCTORS CARE - KNOXVILLE	865-675-3311	101 GLENLEIGH COURT		KNOXVILLE
	edit 4	.011	11514097	EDDY MEDICAL GROUP - DANIELSVILLE	706-795-2211	280 GENERAL DANIELS AVENUE		DANIELSVILLE
	edit 4	012	10844594	ALLS CHURCH URGENT CARE	703-538-1505	920-B WEST BROAD STREET		FALLS CHURCH
	edit 4	013	11527223 (	OUTER BANKS URGENT CARE	252-261-8040	4923 SOUTH CROATAN HIGHWAY		NAGS HEAD
	edit 4	014	11241421	OWNSEND CLINIC	904-461-1901	4475 UNITED STATES HIGHWAY 1		SAINT AUGUSTIN
	edit 4	+015	11521794	MELON IMMEDIATE CARE	434-929-1095	200 AMELON SQUARE		MADISON HEIGH
	edit 4	016	11241568	AIRFAX CONVENIENT CARE	703-849-0900	8301 ARLINGTON BOULEVARD	SUITE 100	FAIRFAX
	edit 4	+017	10193980	RINCETON PRIMARY AND URGENT CARE CENTER LIMITED LIABILITY COMPANY	609-919-0009	707 ALEXANDER ROAD	SUITE 201	PRINCETON
	edit 4	018	10422029	IOLZER CLINIC - JACKSON	740-395-8805	280 PATTONSVILLE ROAD		JACKSON
	edit 4	4019	10993674	MYRNA MEDICAL AID UNIT	302-659-4545	100 SOUTH MAIN STREET		SMYRNA
	edit 4	+020	10425865	ENNESSEE URGENT CARE ASSOCIATES - ANTIOCH FACILITY	615-399-6898	2553 MURFREESBORO PIKE		NASHVILLE
	edit 4	i021	10194150	AINT JOSEPH HEALTH NETWORK URGENT AND DIAGNOSTIC CARE	610-913-1234	45 SOUTH PINE STREET		ELVERSON
	edit 4	4022	11513823	MERGENCY CARE AT LAKE JOY	478-987-0323	1118 STATE HIGHWAY 96		KATHLEEN
	edit 4	1023	10421601		321-728-0000	15 FAST HIBISCUS BOULEVARD		MELBOURNE
	edit 4	1024	10842334	NDIAN RIVER WALK IN CLINIC	772-778-1400	1880 37TH STREET		VFRO BEACH
	edit 4	1025	10421917	ROMED MINOR EMERGENCY CENTER	704-216-2504	628 WEST INNES STREET		SALISBURY
	edit 4	1026	10421908	IFXTCARE LINGENT CARE - LUMBERTON	910-738-7241	2601 NORTH FLM STREET	PROFESSIONAL PLAZA SUITE A	LUMBERTON
	edit 4	4027	10425860	TEMMONS LIRGENT AND PRIMARY CARE	336-712-8225	2245 I FWISVILLE CLEMMONS ROAD	SUITE C	CLEMMONS
	edit 4	1028	10425857		252-808-3696	5059 STATE HIGHWAY 70 WEST		MOREHEAD CITY
	-Page		Whole resu	t - Modify - Selected (0) - Export (4,810)	232-000-5050	SUSS STATE HIGHWAT /U WEST		MORENLAD CIT
	12345	97	4.810 rot					



```
Data Storage and Management
Database Choice: PostgreSQL / PostGIS
Database Interfaces: Adminer, Rstudio, Jupyter, psql
```

Very fast access to geo files and database GIS functions

```
# create db connection
con <- sdalr::con_db(dbname = "sdad", host = "127.0.0.1", port = 5433, user = "anonymous", pass =
"anonymous")
# create SQL query
sql <- "SELECT distinct \"GEOID10\" geoid, geometry
FROM tl_2018_19_tabblock10 where left(\"GEOID10\", 5) = '19127'''
# get census blocks
marshall_county_blocks <- sf::st_read(con, query = sql) %>%
st transform(crs = 4269)
```



### Code Management: GitLab Multiple Branch Management Critical in Team Science

GitLab.org	/ GitLab Community Editio	on 🗸 🦊		This project Search	् 📠 🌒 -
	Project	t Activity Repository Pipelines Graphs Issues 5	,483 Merge Requests 413	Snippets	Ø - ×
		Pipelines Builds Environments (	Cycle Analytics		
All 31,341	Running 63 Branches	Tags			Run pipeline CI Lint
Status	Pipeline	Commit	Stages		
Image: Provide the second s	#5233451 by 🌒 🚺 🔒	<b>20916-issues-mrs-linel</b> → e3612f86  Add content_class for limited width	<ul> <li>✓ (□) (◎)</li> </ul>		± - ×
Image: Provide the second s	#5233393 by 🏶 🚺 latest	P fix/refactor-cycle-ana → ab98279e big refactor based on MR feedback	© (I) (I)		± - ×
Image: Provide the second s	#5233299 by 🍣 🚺 latest	<ul> <li>P dz-nested-groups -&gt; 2b2b39f0</li> <li>         Add support for nested namespaces in th     </li> </ul>	<ul> <li>○ (0) (0)</li> </ul>		± - ×
Image: Provide the second s	#5233284 by 🎯 🛛 latest	P master ~ 2f91c0ee Ø Merge branch 'cleanup-common_utils.js'	⊘ (0)-(◎)-(◎)		± - ×
Image: Provide the second s	#5233164 by 🜒 🚺 latest	P 25106-hide-issue-mr-bu → b3d15756     adds changelog			± - C ×
S canceled €	#5233032 by 🧐	P 25106-hide-issue-mr-bu → 2fedc86e adds check for logged in user in group iss	<b>⊗</b> -⊙-→ <sup>8</sup>	00:00 52 minutes ago	C
Image: Provide the second s	#5232996 by 🧐 🚺 latest	<ul> <li>P 25209-improve-length-v → 62d5439e</li> <li>Wse :maximum instead of :within for leng</li> </ul>	© () ()		± - C ×
	-	γ <b>master</b> → af19d56f			

### Data Profiling & Preparation

#### Preferred Platform: RStudio Server

- ματα\$

#### Git Integration

R File Edit Code View Plots Session Build Debug Profile Tools Help				aschroed 🕞 🛛 😃
💙 🛨 - 🐑 📹 - 🗐 着 🍺 Go to file/function 👘 🐇 - 🗮 - Addins -				🔋 sdad-data 👻
🗈 census lodes.R 🛪 🔊 census acs vulnerability index.R 🛪 🖻 gtfs.R 🛪 🔊 api testing.R X 🖏 get.bg. gravity.R X	Environr	nent History Connections Git		
🖛 🗈 🖉 Source on Save 🔍 🎘 📲	ırce - 🗐 📄 Diff	📝 Commit   🖊 Pull   🏠 Push   🕑   🌞 🕤		Tanta and the second se
1 library(data.table)	Staged S	Status Path		
2 library(ggplot2)		M api_testing.R		
4 source("functions/get lodes R")		<pre>M functions/get_bg_gravity.R</pre>		
5 source("functions/get_bg_growity.R")		Sources/transit/gtts.k		
6 source("functions/theme_map.R")				
7 8 con $x_{-}$ stall $x_{-}$ stall $x_{-}$ stall host $-$ "127 0 0 1" nont $-$ 5433 user $-$ "approximate" pass $-$ "approximate"				
9				
10 # get census block group geographies				
11 sql <- "SELECT distinct \"GEOID\" gooid, geometry				
12 FROM $t_2dis_{19}$ go where $lett(('dellow'', 5) = '19127''''''''''''''''''''''''''''''''''$				
14				
15 # exclude certain block groups				
16 excld <- c("191279501004","191279502002","191279502003","191279502001","191279503002","191279503003","191279503004","191279504004") 17 ho gens <- ho gens [] ho gens [] ho gens <- ho gens <- ho gens <- ho gens ]] ho gens <- ho gens <- ho gens <- ho gens [] ho gens <- ho ge	Files	Plots Packages Help Viewer		
18	ta New	Folder 🛛 🖓 Upload 🛛 🕸 Delete 🏹 Rename 🗌	🍄 More 👻	5
19 # Get LODES job count data		Home > git > sdad-data > functions		
20 lodes_ia_2015 <- data.table::setDT(read_lodes("ia", "od", "aux", "JT00", "2015", "data/sdad_data/original/CENSUS/LODES")) 21 lodes_ia_2015 <- data.table::setDT(read_lodes("ia", "od", "aux", "JT00", "2015", "data/sdad_data/original/CENSUS/LODES"))		▲ Name	176	Modified
21 todes_td_2015[, w_geocode := ds.thuracter(w_geocode)] 22				
23 # get gravity indexes by block group		cartogran .c.R	721 B	M. 22, 2019, 9:52 AM
<pre>24 gravity_idx &lt;- get_bg_gravity(bg_geos\$geoid, block_counts_df = lodes_ia_2015, block_geoid = "w_geocode", block_cnt = "\$000")</pre>		datenfo.R	453 B	Mar 22, 119, 9:52 AM
$25$ gravity_laX[, bglax_lg := log(bglaX)] $26$ gravity_idx(haidy lg <= scale(orgaity, idx(haidy lg center=min(organity, idx(haidy lg) scale=diff(range(orgaity, idx(haidy lg)))		o_names.R	2 KB	Mar 22, 201 9:52 AM
		degrees2meters.R	ОВ	Mar 22, 2019, 572 AM
<pre>28 gravity_idx[, rank := cut(bgidx_lg,breaks=quantile(bgidx_lg,probs=seq(0,1,by=0.2)),labels=1:5,include.lowest=TRUE)]</pre>		FCClocation2FIPS.R	2.2 KB	Mar 22, 2019, 9:5. AM
29 30 # menon amounty indexes with block amoun apparentias	· · · · · · · · · · · · · · · · · · ·	fread_combine.R	352 B	Mar 22, 2019, 9:52 💦 1
31 tomap <- merae(arayty_idx, ba_aeos, by = "aeoid")	· · · · · · · · · · · · · · · · · · ·	get_acs.R	4.9 KB	Mar 22, 2019, 9:52 AN
32 # convert to sf		get_bg_gravity.R	3.7 KB	Mar 26, 2019, 4:54 AM
33 tomap_sf <- ofus_st(tomap)		get_cont_geo_id.R	431 B	Mar 22, 2019, 9:52 AM
21:1 (Trianel) :	R Script	get_kidscount.R	3.1 KB	Mar 22, 2019, 9:52 AM
Cop c Terminal ×		get_lodes.R	3 KB	Mar 22, 2019, 9:52 AM
Terminal 1 v aschroed@docker-s-12vcpu-48gb-nyc1-01: ~/git/sdad-data		get_osm.R	1.6 KB	Mar 22, 2019, 9:52 AM
drwxr-xr-x 2 aschroed aschroed 4.0K Mar 22 14:52 pres-figure		get_webpage_links.R	117 B	Mar 22, 2019, 9:52 AM
rw-rr 1 aschroed aschroed 893 Mar 22 14:52 pres.Rpres		latlong2county.R	591 B	Mar 22, 2019, 9:52 AM
-rw-rr 1 aschroed aschroed 1.2K Mar 22 14:52 pres.md	📕 🗖 🖻	metadata.R	5.3 KB	Mar 22, 2019, 9:52 AM
-rw-rr- 1 aschroed aschroed 205 Mar 26 08:00 sdad-data.Rproj		normalize_colname.R	781 B	Mar 22, 2019, 9:52 A
drwxr-xr-x 11 aschroed aschroed 4.0K Mar 22 14:52 sources	Dia d	recreate_db.R	1.7 KB	Mar 22, 201 <u>9, 9:52</u> M
-rw-rr 1 aschroed aschroed 3.9M Mar 22 14:52 temp.RDS		theme map.R	840 B	Mar 22, 2019, 9: AM
Arr-r 1 aschroed aschroed 8.0M Mar 22 14.32 temp. Kouta		zip2fips_create.R	625 B	Mar 22, 2019 52 AM
-rwr- 1 aschroed aschroed 3.9M Mar 22 14:52 tempz.RData			0250	
-rw-r, 1 aschroed aschroed 1.2K Mar 22 14:52 test.csv				
-rw-r				

aschroed@docker-s-12vcpu-48gp-nyc

ERSITY GINIA

Server Files

### Data Profiling: Quality

#### **Completeness**

percentage of elements properly populated

e.g. Testing for NULLs and empty strings where not appropriate

#### Value Validity

percentage of elements whose attributes possess meaningful values e.g. A comparison constraint like {male; female} or an interval constraint like age = [0,110]

#### **Consistency**

a measure of the degree to which two or more data attributes satisfy a well-defined dependency constraint – relationship validation

e.g. Zip-code – state consistency or gender – pregnancy consistency <u>Uniqueness</u>

the number of unique values taken by an attribute, or a combination of attributes in a dataset

e.g. Frequency distribution of an element

note. The more homogeneous the data values of an element, the less useful the element is for analysis

#### **Duplication**

a measure of the degree of replication of distinct observations per observation unit type

e.g. Greater than 1 registration per student per official reporting period note. Duplication can occur as a result of choice of level of aggregation; for example, aggregating to a single student registration per academic year when registration information is actually collected multiple times per academic year

#### Race Type Report

#### SUMMARY OF DATA ELEMENT

Race Type is recorded for each student as the one or more races they identify with. This set was collected from students in the school system during the school years of 2005-2015. During our collection period, starting with the 2010-2011 school year, the race type definition expanded to allow students to choose multiple races and separately specify Hispanic ethnicity. In this process categories were eliminated. Eliminated categories included the 'Unspecified or Unknown' category and the 'Hispanic' category. A student's Hispanic ethnicity is now captured under an ethnic flag.

#### Number of Unique Values: 8

#### value value\_description

- 99 Multiple race types reported1 American Indian or Alaskan Native
- 2 Asian
- Black or African American
- Hispanic (valid before 2010 school year)
- 5 White
- 6 Native Hawaiian or Other Pacific
- Islander 0 Unspecified or Unknown (valid before 2010 school year)



#### DATA ANALYTICS SUMMARY OF RESULTS

Test	Measurementvalues	Value
Completeness	Number of missing values	0
	Percent of complete values	100 %
Validity	Number of invalid values	0
	Percent of valid values	100 %
Uniqueness	Number of unique values	8
Record Consistency I	Number of inconsistent records	0
	Percent consistent records	100 %
Record Consistency II	Number of inconsistent records	2830444
	Percent consistent records	93.28 %

### Data Profiling: Valid Values

- Data elements with proper values have value validity
- The percentage of data elements whose attributes possess values within the range expected for a legitimate entry is a measure of value validity
- Checking for value validity generally comes in the form of straight-forward domain constraint rules
  - How many entries contain non-valid values for a non-empty text field representing gender?
    - < count gender where gender is not (male, female) >
  - How many entries contain non-valid values for a non-empty integer field representing age?
    - < count age where age is not between [0, 110] >

#### Pulled from current James City County MLS Data

zip_code	area	subdivision	neighborhood	zoning	parcel_id
23185	JCC	Governors Land	River Reach	R-4	4511000022
23188	JCC	Wellington	/	RESIDENT	1330800178
23188	JCC	Powhatan Secondary		RES	3741600013
23185	JCC	Kingsmill	Padgetts Ordinary	R 4	5041100213
23185	JCC	Pointe @ Jamestown	(	RES	4640600108
23185	JCC	Paddock Green	Paddock Green	R1	

Comparison constraint: **zoning 2015, James City County**= {A-1, R-1, R-2, R-3, R-4, R-5, R-6, R-7, R-8, LB, B-1, M-1, M-2, RT, PUD, MU, PL, EO}

- During Data Profiling issues are described, not "fixed"
- The appropriate fix depends upon the needs of the research
- It may be appropriate to simply normalize all zoning entries to the five major categories of zoning: Residential, Mixed Residential-Commercial, Commercial, Industrial, and Special



### Data Profiling: Consistency

#### RECORD CONSISTENCY II

Find Records with an inconsistent relationship between Race Type and School Year . Check if there any records that have a race\_type of '4' after the 2009 school year.

Number Inconsistent: 2830444

Percent Consistent: 93.28 %

**Record Consistency** The concept of record consistency is best understood as the degree of logical agreement "between" record field values in either a single dataset or between two or more datasets. Simple Example: location disagreement between zip code and state FIPS code.

**Longitudinal Consistency** An inconsistency in the data when checked over time (longitudinally), to see if the same value is recorded for every new record when it should be (i.e. birthdate and other demographics).



## Data Profiling & Preparation

The **Data Preparation** Phase includes the activities necessary to **"fix"** the issues of Quality, Structure, and Metadata discovered during Data Profiling – activities can include: Cleansing

Missing Values Date Formats Nominal => numeric Outliers

Inconsistent Data

**De-duplication** 

Transformation

Aggregation

Normalization

Smoothing/Winsorization

Imputation

Feature Construction

Restructuring

#### Current Structure of Williamsburg MLS Data

List Number	Agency Name	Agency Phone	Agency Email	Listing Agent	Listing Agent Phone	Listing Agent Email	Co-Listing Agent	Property Type	Card Format
Book Section	Selling Agency	Selling Agency Phone	Selling Agency Email	Selling Agent	Selling Agent Phone	Selling Agent Email	Co-Selling Agent	End Date	book_sec
Listing Date	Sold Date	Under Cont. Date	Fall-thru Date	Status	Status Change	Withdraw Date	Cancel Date	Contingent	Cont. Remarks
Orig. List Price	Price	Sold Price	high_price	Low Price	assessed_val	Partial Tax Assmnt	financing	Area	Relocation
St. #	box_nbr	St. Dir.	Street Name	Address 2	streetdirsuffix	Street Suffix	carrier_route	City	State
county	country	Zip Code	geo_county	Taxes	geo_lat	geo_lon	Est. Fin. SqFt	sqft1	sqft2
sqft3	sqft4	Year Built	2+ Bdroms on 1st Flr	Realtor.com Type	lot_size	Total Acres	Condo Level	sell_broker_comm	Variable Commission
stories	Total Rooms	Total Bedrooms	total_bath	Baths - Full	Baths - Half	baths_3_4	Garage Type	garage_stall	Water Frontage
Zoning	taxes	Tax Year	Subdivision	Public Remarks	Agent Remarks	Parcel ID	Legal Description	Directions	Foreclosure
Owner Phone	Owner Name	Neighborhood	mod_timestamp	Ltd Service Agent	Occupied By	Owner/Agent	Mster Bdrm 1st Floor	SqFt Source	Listing Type
# Stories	# Fireplaces	Golf Frontage	IDX Y/N	Supplement Attached	Seller Concession(s)	Special Assmnts	Туре	Rollback Taxes	userdefined16
SellingBroker Incent	Ownership	Describe Concession	How Sold	Selling Broker Comp	userdefined22	Assessed Value	Est.Unfinished Sq Ft	Tax Rate	Garage Bays
userdefined27	userdefined2 8	userdefined29	userdefined30	Est. Closing Date	userdefined32	userdefined33	Lot Description	Short/CompromiseSa le	userdefined36
userdefined37	userdefined3 8	userdefined39	userdefined40	userdefined41	userdefined42	userdefined43	userdefined44	userdefined45	userdefined46
userdefined47	userdefined4 8	userdefined49	userdefined50	userdefined51	userdefined52	userdefined53	userdefined54	userdefined55	userdefined56
Photo URL	Days on Market	Rooms	Features						



### Recent Example: Distinct Counts of Children 0-5 Datasets Under Consideration

Finding the best representations of Who, What, When & Where Across Agencies, Across Datasets

Source	Datasat	Demographics	Ethnicity	Conder	A.m.	Location	1 Sanutra	Time
VDSS	Customers By Year	cust_race_is_amer_indian_alaska_native_ind customer_race_is_asian_indicator customer_race_is_black_indicator customer_race_is_hawaiian_pacific_islander_ind customer_race_is_other_indicator customer_race_is_white_indicator	ethnicity_code	gender_code	age_class_code age_group_code age_type_code month_of_birth year_of_birth	-	-	calendar_year_number service_year
VDSS	SNAP Customers By Year	-		-		zip_code	county_fips_code	calendar_year_number
VDSS	TANF Customers By Year			-		zip_code	county_fips_code	calendar_year_number
VDSS	Foster Customers By Year	-		-			county_fips_code	calendar_year_number
VOCS	OCS Services By Year		-	-			-	service_begin_date service_end_date service_duration program_year
VDOE	Unique Student Listing	race_type		gender	birth_month birth_year [grade_code]		[division_number_reporting_school_numbe]	school_year entry_date
VDOE	VPI+	students_race_ethnicity	students_race_ethnicity	gender	birth_month birth_year [grade]		[school]	assessment_date school_year



### Recent Example: Distinct Counts of Children 0-5 Longitudinal Consistency

• Each dataset profiled to discover the most consistent recording of demographic information over time





### Recent Example: Distinct Counts of Children 0-5 Transformation & Linkage Enables Distinct Count Cross-Tabulation

Distinct Count Cross-Tabulation Service By Race/Ethnicity By Year											
Services	Demographic	2013	2014	2015	2016						
snap & tanf	black hispanic or latino not reported	1159	1710	1766	1333						
snap & tanf	black not hispanic or latino	72710	62084	56655	40908						
snap & tanf	other hispanic or latino	2112	1752	1529	1025						
snap & tanf	other hispanic or latino not reported	143	162	129	66						
snap & tanf	other not hispanic or latino	1870	1616	1287	877						
snap & tanf	race not reported hispanic or latino	1438	1123	1064	749						
snap & tanf	race not reported hispanic or latino not reported	4461	5727	8795	6941						
snap & tanf	race not reported not hispanic or latino	5409	4375	4190	3213						
snap & tanf	white hispanic or latino	5357	4412	4325	3424						
snap & tanf	white hispanic or latino not reported	902	1889	1958	1264						
snap & tanf	white not hispanic or latino	41489	34649	30686	21683						



### Recent Example: Distinct Counts of Children 0-5 Distinct Count Cross-Tabulation Enables Distinct Count Plots





Recent Example: Distinct Counts of Children 0-5 Distinct Count Cross-Tabulation Enables Creation of New Indices

 Example Composite Index Combining SNAP (Nutrition) and OCS (Behavioral Assistance) Data





### Data Analysis & Data Product Creation

#### Preferred Platform: RStudio Server

-uutas

#### Git Integration

Rile Edit Code View Plots Session Build Debug Profile Tools Help				aschroed 🗈 🕚	5
🔍 🕶 📲 - 🧐 📹 - 📰 📥 👘 Go to file/function 🔤 😓 - 🧮 - Addins -				🔋 sdad-data 🕙	-
🗊 census lodes.R 🛪 🔊 census acs vulnerability index.R 🛪 🔊 otfs.R 🛪 🔊 api testino.R 🛪 🔊 oet bo oravity.R 🛪	_	Environment History Connections Git			
A Source on Save	🖸 🔄 🔁 Source 🚽 🗐	🗧 Diff 🔽 Commit 👎 Pull 🔶 Push 🕒 🔮	a -		
1 library(data.table)		Staged Status Path			
2 library(ggplot2)		M api_testing.R			
3 library(sf)		functions/get_bg_gravity.R			
<pre>4 source('tunctions/get_loaes.k') 5 source('functions/get_loaes.k')</pre>		sources/transit/gtfs.R			
6 source("functions/theme_map.R")					
7					
8 con <- sdalr::con_db(dbname = "sdad", host = "127.0.0.1", port = 5433, user = "anonymous", pass = "anonymous")					
10 # aet census block aroup aeoaraphies					
11 sql <- "SELECT distinct \"GEOID\" geoid, geometry					
12 FROM tl_2018_19_bg where left(\"GEOID\", 5) = '19127'"					
13 bg_geos <- sf::st_read(con, query = sql)					
15 # exclude certain block aroups					
16 excld <- c("191279501004", "191279502002", "191279502003", "191279502001", "191279503002", "191279503003", "191279503004", "191279504004")		Files Plots Packages Help Viewer			
17 bg_geos <- bg_geos[!bg_geos\$geoid %in% excld,]		📬 New Folder 🛛 🖓 Upload 🚱 Delete 📑 Rename	e   🏠 More 🗸 🔰	C	
18 19 # Cat LOBES job count data		$\square$			
20 Jobes in 2015 <- data.table::setDT(read lodes("ia", "od", "aux", "JT00", "2015", "data/sdad data/original/CENSUS/LODES"))		▲ Name	Size	Modified	
21 lodes_ia_2015[, w_geocode := as.character(w_geocode)]		<b>t</b>			
22		Cartogram.c.R	721 B	M. 22, 2019, 9:52 AM	
23 # get gravity indexes by block group 24 gravity indexes by block group 24 gravity indexes by a gravity indexes and index counts of $-1$ loops in 2015 block gravity indexes block gravity indexes by block gravity inde		■ <sup>®</sup> dat of o.B	453 B	Mar 22, 019, 9:52 AM	
25 gravity_tax[, baidx_la := loa(baidx]]		D names R	2 KR	Mar 22, 20, 9.52 AM	
26 gravity_idx\$bgidx_lg <- scale(gravity_idx\$bgidx_lg,center=min(gravity_idx\$bgidx_lg),scale=diff(range(gravity_idx\$bgidx_lg)))		degrees2meters P	0 R	Mar 22, 2019 52 AM	
			2240	Mar 22, 2019, 512 AM	
28 gravity_laxL, rank := cut(bglax_lg,breaks=quantile(bglax_lg,pross=seq(0,1,by=0.2)),labels=1:5,lnclude.lowest=1kUE)] 29		E freed combine D	2.2 KB	Mar 22, 2019, 9.5. 4M	
30 # merge gravity indexes with block group geographies			352 B	Mar 22, 2019, 9:52 An	
31 tomap <- merge(gravity_idx, bg_geos, by = "geoid")		get_acs.R	4.9 KB	Mar 22, 2019, 9:52 AN	
32 # convert to sf		get_bg_gravity.R	3.7 KB	Mar 26, 2019, 4:54 AM	
20 comp_sr csr(comp)		get_cont_geo_id.R	431 B	Mar 22, 2019, 9:52 AM	
21:1 (Frel) :	R Script	get_kidscount.R	3.1 KB	Mar 22, 2019, 9:52 AM	
Contract Terminal ×	-6	get_lodes.R	3 KB	Mar 22, 2019, 9:52 AM	
Terminal 1 +   aschroed@docker-s-12vcpu-48gb-nyc1-01: ~/git/sdad-data	4	🔲 🖻 get_osm.R	1.6 KB	Mar 22, 2019, 9:52 AM	
drwxr-xr-x 2 aschroed aschroed 4.0K Mar 22 14:52 pres-figure		🔲 🖻 get_webpage_links.R	117 B	Mar 22, 2019, 9:52 AM	
-rw-r-r 1 aschroed aschroed 893 Mar 22 14:52 pres. Rpres		Iatlong2county.R	591 B	Mar 22, 2019, 9:52 AM	
-rw-rr-1 aschroed aschroed 1.2K Mar 22 14:52 pres.ma		metadata.R	5.3 KB	Mar 22, 2019, 9:52 AM	
-rw-rr- 1 aschroed aschroed 205 Mar 26 08:00 sdad-data.Rproj		normalize_colname.R	781 B	Mar 22, 2019, 9:52 A	
drwxr-xr-x 11 aschroed aschroed 4.0K Mar 22 14:52 sources		Precreate db.R	1.7 KB	Mar 22, 2019, 9:52 M	
-rw-rr 1 aschroed aschroed 3.9M Mar 22 14:52 temp.RDS		theme map.R	840 B	Mar 22, 2019, 9: AM	
The re-re-1 aschoed as		zip2fips_create_R	625 B	Mar 22, 2019 52 AM	
-rwr 1 aschroed aschroed 3.9M Mar 22 14:52 tempz.RData			0250	Mul 22, 2015	
-rw-r 1 aschroed aschroed 1.2K Mar 22 14:52 test.csv					
-rw-rr- 1 asch					Ľ

aschroed@docker-s-12vcpu-48gp-nyc

ERSITY GINIA

Server Files

### Data Ingestion Marshalltown

#### Shapefiles from our GIS database (PostGIS)

```
# create db connection
con <- sdalr::con_db(dbname = "sdad", host = "127.0.0.1", port = 5433, user = "anonymous", pass =
"anonymous")
# create SQL query
sql <- "SELECT distinct \"GEOID10\" geoid, geometry
FROM tl_2018_19_tabblock10 where left(\"GEOID10\", 5) = '19127'''
# get census blocks
marshall_county_blocks <- sf::st_read(con, query = sql) %>%
st transform(crs = 4269)
```



### Data Ingestion Marshalltown

#### library(tidytransit)

# get Marshalltown feed URL feed\_url <- feedlist\_df %>% setDT(.) %>% .[loc\_t %like% "Marshalltown, IA", url\_i]

# read gtfs data from url
gtfs <- read\_gtfs(feed\_url, geometry = TRUE, frequency =
TRUE)</pre>

Transit route data from transitfeeds.com

		- 71	
) g	tfs	list [28] (S3: gtfs)	List of length 28
	agency	list [1 x 7] (S3: spec_tbl_df, tb	A tibble with 1 rows and 7 columns
	areas	list [0 x 2] (S3: data.table, dat	A data.frame with 0 rows and 2 columns
	calendar_attributes	list [4 x 2] (S3: data.table, dat	A data.frame with 4 rows and 2 columns
	calendar_dates	list [34 x 4] (S3: spec_tbl_df, 1	A tibble with 34 rows and 4 columns
	calendar	list [4 x 11] (S3: spec_tbl_df, 1	A tibble with 4 rows and 11 columns
	directions	list [0 x 3] (S3: data.table, dat	A data.frame with 0 rows and 3 columns
	fare_attributes	list [1 x 7] (S3: spec_tbl_df, tb	A tibble with 1 rows and 7 columns
	fare_rider_categories	list [0 x 3] (S3: data.table, dat	A data.frame with 0 rows and 3 columns
	fare_rules	list [11 x 5] (S3: spec_tbl_df, 1	A tibble with 11 rows and 5 columns
	farezone_attributes	list [0 x 2] (S3: data.table, dat	A data.frame with 0 rows and 2 columns
	feed_info	list [1 x 10] (S3: spec_tbl_df, 1	A tibble with 1 rows and 10 columns
	frequencies	list [0 x 5] (S3: spec_tbl_df, tb	A tibble with 0 rows and 5 columns
	linked_datasets	list [0 x 7] (S3: data.table, dat	A data.frame with 0 rows and 7 columns
	rider_categories	list [0 x 2] (S3: data.table, dat	A data.frame with 0 rows and 2 columns
	routes	list [11 x 12] (S3: spec_tbl_df,	A tibble with 11 rows and 12 columns
	runcut	list [0 x 9] (S3: data.table, dat	A data.frame with 0 rows and 9 columns
	shapes	list [2174 x 5] (S3: spec_tbl_d	A tibble with 2174 rows and 5 columns
	stop_attributes	list [13 x 2] (S3: data.table, da	A data.frame with 13 rows and 2 columns
	stop_times	list [792 x 22] (S3: spec_tbl_d	A tibble with 792 rows and 22 columns
	stops	list [76 x 15] (S3: spec_tbl_df,	A tibble with 76 rows and 15 columns
	timetable_stop_order	list [0 x 5] (S3: data.table, dat	A data.frame with 0 rows and 5 columns
	timetables	list [0 x 16] (S3: data.table, da	A data.frame with 0 rows and 16 columns
	transfers	list [1 x 4] (S3: spec_tbl_df, tb	A tibble with 1 rows and 4 columns
	trips	list [58 x 18] (S3: spec_tbl_df,	A tibble with 58 rows and 18 columns
	stops_sf	list [76 x 14] (S3: sf, tbl_df, tł	A tibble with 76 rows and 14 columns
	routes_sf	list [11 x 2] (S3: sf, tbl_df, tbl	A tibble with 11 rows and 2 columns
	stops_frequency	list [169 x 6] (S3: data.table, (	A data frame with 169 rows and 6 columns
	routes_frequency	list [11 x 5] (S3: tbl_df, tbl, da	A tibble with 11 rows and 5 columns



### Data Ingestion Marshalltown

acs\_vars <- c( "B25070\_001","B25070\_010", "B25091\_001","B25091\_011","B25091\_022", "B25044\_001","B25044\_003","B25044\_010", "B22010\_001","B22010\_002", "B17021\_001","B17021\_002"

acs\_est <- get\_acs(geography="block
group",state=state\_names,county=county\_names,</pre>

```
variables=acs_vars,year=year,cache_table=TRUE,out
put="wide", geometry = TRUE)
```

Vulnerability data composited from Census ACS housing, transportation, and nutrition subsidy variables

NAME ÷	B25070_001E <sup>‡</sup>	B25070_001M <sup>‡</sup>	B25070_010E <sup>‡</sup>	B25070_010M <sup>‡</sup> B2
Block Group 1, Census Tract 9501, Marshall County, I	24	19	3	4
Block Group 2, Census Tract 9501, Marshall County, I	16	12	0	9
Block Group 3, Census Tract 9501, Marshall County, I	56	38	3	5
Block Group 4, Census Tract 9501, Marshall County, I	15	8	4	4
Block Group 1, Census Tract 9502, Marshall County, I	76	32	3	5
Block Group 2, Census Tract 9502, Marshall County, I	83	55	2	5
Block Group 3, Census Tract 9502, Marshall County, I	91	28	5	7
Block Group 1, Census Tract 9503, Marshall County, I	45	26	0	9
Block Group 2, Census Tract 9503, Marshall County, I	76	27	10	12
Block Group 3, Census Tract 9503, Marshall County, I	28	17	3	4
Block Group 4, Census Tract 9503, Marshall County, I	87	43	6	6
Block Group 1, Census Tract 9504, Marshall County, I	90	29	2	3
Block Group 2, Census Tract 9504, Marshall County, I	18	17	0	9
Block Group 3, Census Tract 9504, Marshall County, I	29	20	3	6
Block Group 4, Census Tract 9504, Marshall County, I	47	27	2	3
Block Group 1, Census Tract 9505, Marshall County, I	288	99	51	45
Block Group 2, Census Tract 9505, Marshall County, I	175	77	38	45
Block Group 3, Census Tract 9505, Marshall County, I	160	81	0	9
Block Group 4, Census Tract 9505, Marshall County, I	140	71	14	17
Block Group 1, Census Tract 9506, Marshall County, I	161	79	5	9
Block Group 2, Census Tract 9506, Marshall County, I	135	63	68	53
Block Group 3, Census Tract 9506, Marshall County, I	315	95	35	25
Block Group 4, Census Tract 9506, Marshall County, I	48	40	11	18



### Gravity Models for Job Access and Transit Access

🖻 cen	sus_lodes.R 🗴 🖻 census_acs_vulnerability_index.R 🗴 🖻 gtfs.R 🗴 🖻 api_testing.R 🗴 🖻 get_bg_gravity.R 🗶	
<b>( )</b>	🛲 📲 🔳 Source on Save 🔍 🎢 🖌 📕	
37	#* @apiTitle Block Group Gravity Index	
38		
39	#* Create a Gravity Model Index using a list of block group geoids (12 characters) and	n
40	#* a data.frame of block geoids and a count of something for each one	$r = \sum_{i=1}^{n} p_i$
41	#* @param bg_geoid_list list of Lensus block group geoids (12 characters)	$E \equiv \sum_{r=1}^{\infty} \frac{1}{r^2}$
42	#* @param block_counts_at data.trame with a geola column (15 characters) and a count of somethic #* @param block_counts_at the name of the column bolding the counts in block counts of	$\frac{1}{i=1}$
44	#* dist mi the distance from each block aroun centernoint from which block will be retrieved	Equation 3: Employment Access Index Definition
45	#* @post /aet_ba_aravity	where:
46	<pre>get_bg_gravity &lt;- function(bg_geoid_list, block_counts_df, block_geoid = "geoid", block_cnt = "</pre>	E is the Employment Access for a given Census block group
47	library(data.table)	n is the total number of Census blocks
48	<pre>for (bgid in bg_geoid_list) {</pre>	p <sub>i</sub> is the number of jobs in the i <sup>th</sup> Census block
49	# Get Blocks	r <sub>i</sub> is the distance (in miles) from the center of the given Census block group to the center of the i''
50	bg_blocks <- get_blocks_within_distance_of_bg(bg = bgid, dist_mi = dist_miles)	Census block
51	# Monao Riocks and LODES	
53	bal count <- merge(ba blocks, block counts df, by $x = "aeoid block", by y = block aeoid)$	
54	setDT(ba1_count)	
55		
56	# Aggregate Jobs per Block	
57	<pre>bg1_count_2 &lt;- bg1_count[,.(block_cnt=sum(as.numeric(get(block_cnt)))), .(geoid_bg, geoid_b</pre>	lock, dist_mi)]
58		
59	# Calculate Index	
60	<pre>idx &lt;- bgl_count_2[, .(block_cnt, a_sqr=dist_mi^2, e=block_cnt/(dist_mi^2))][a_sqr==0, e := idv dt : data table(accid</pre>	0_L, sum(e)_
62	lax_at <- aata.table(geola = bgla, bglax = lax)	
63	# Combine	
64	<pre>if (exists("idxes") == TRUE) idxes &lt;- rbindlist(list(idxes, idx_dt))</pre>	
65	else idxes <- idx_dt	
66	}	
67	idxes	
68	}	
69		

\*

ッソ/IRG]

### Data Visualization & Sharing: Maps (ggplot)

Combining Administrative and Survey Data to Relate Job Availability, Economic Vulnerability and Access to Transit



istics (LODES),

General Transit Feed Spe





### Data Visualization & Sharing Interactive Maps (leaflet)





### Data Visualization & Sharing Interactive Maps (leaflet)





## Data Visualization & Sharing Dashboards (R Shiny, Django)



Clarendon has over 40 restaurants (pinned in blue circles) with ABC (Alcohol and Beverage Control) licenses and an average of 5,500 patrons per weekend night. Each year, approximately 580,000 patrons visit Clarendon between 21:00 and 03:00, especially during holidays and special "drinking" events. Alcoholletated crime counts in Clarendon on Friday. Saturday, and Sunday between 21:00 and 03:00 for 2015-2017 are given in yellow and green circles.



					1							
M				_ Crime Type								
		DUI		-								
۵	1											
											-	Arlington Crime Data
		Copy	/ CSV	Excel Print						5	Search:	
			id 🍦	description 🔶	location 🕴	latitude 🍦	longitude 🔶	start 🍦	end 🍦	year 🔷	month	day_of_week
		1	2018- 05310007	DUI	S WALTER REED DR / 18TH ST S	38.85299031	-77.08812649	2018-05- 31T00:43:00Z	2018-05- 31T00:43:00Z	2018	5	Thursday
		2	2018- 05280152	DUI 3+ OFFENSE OR 2+ FELONY OFFENSE	1XX N GLEBE RD	38.87262371	-77.10374707	2018-05- 28T16:51:00Z	2018-05- 28T16:51:00Z	2018	5	Monday
		3	2018- 05280057	DUI	N LYNN ST / LEE HWY	38.89716894	-77.06996344	2018-05- 28T05:00:00Z	2018-05- 28T05:00:00Z	2018	5	Monday
		4	2018- 05270039	DUI	ARLINGTON BLVD / N COLUMBUS ST	38.86547337	-77.11670666	2018-05- 27T04:04:00Z	2018-05- 27T04:04:00Z	2018	5	Sunday
		5	2018- 05260260	DUI	XX 8468436440000000	38.84758519	-77.08140523	2018-05- 26T23:40:00Z	2018-05- 26T23:40:00Z	2018	5	Saturday
		6	2018- 05260243	DUI	13TH ST S / S GEORGE MASON DR	38.85768855	-77.09867447	2018-05- 26T22:37:00Z	2018-05- 26T22:37:00Z	2018	5	Saturday
		7	2018- 05260178	DUI	23XX 25TH ST S	38.84887425	-77.07555001	2018-05- 26T17:15:00Z	2018-05- 26T17:30:00Z	2018	5	Saturday
		8	2018- 05260025	DUI	25XX S WALTER REED DR	38.84614347	-77.10075429	2018-05- 26T01:30:00Z	2018-05- 26T01:30:00Z	2018	5	Saturday
		9	2018- 05250034	DUI	N HIGHLAND ST / 9TH RD N	38.88215652	-77.09260151	2018-05- 25T03:08:00Z	2018-05- 25T03:08:00Z	2018	5	Friday
		10	2018- 05250025	DUI	10TH ST N / FAIRFAX DR	38.88351863	-77.09836161	2018-05- 25T02:04:00Z	2018-05- 25T02:06:00Z	2018	5	Friday

JNIVERSITY VIRGINIA



VERSITY IRGINIA



"JNIVERSITY "VIRGINIA

### APIs (Plumber, OpenCPU)

- 1. Distance from BG center-points to Block center-points
- 2. Gravity model API

← → C ① Not Secure   sdad.policy-analytics.net:8000/get_blocks_within_distant	nce_of_bg?bg=510131016033&dist_mi=1.25 🔹 🖈 🙀 😰 🖉 🖉 🗐 🐐 🖕 🔍 🔘
👯 Apps 🚔 Log in   Biocompl 🗎 Programming 🗎 SysAdmin 🗎 SDAL 🗎 Data 🗅	VLDS Data Quality 📄 Data Science 📄 Data Policy 📧 Mapping Informati 👛 Files - OneDrive 👩 Google Keep 🔤 Verizon Messages 🌀 Android Device M 🗎 Boats 🗎 Math 🗎 A
[{"geoid_bg":"510131016033","geoid_block":"510131018021004","dist_mi":1.164}	<pre>,{"geoid_bg":"510131016033","geoid_block":"510131018021008","dist_mi":1.1608},{"geoid_bg":"510131016033","geoid_block":"510131018021003","dist_mi":1.0969},</pre>
{"geoid_bg":"510131016033","geoid_block":"510131018023002","dist_mi":1.2143}	,{"geoid_bg":"510131016033","geoid_block":"510131018021005","dist_mi":1.231},{"geoid_bg":"510131016033","geoid_block":"510131034011002","dist_mi":1.0212},
{"geoid_bg":"510131016033","geoid_block":"510131034011011","dist_mi":1.1844}	,{"geoid_bg":"510131016033","geoid_block":"510131034011013","dist_mi":1.0252},("geoid_bg":"510131016033","geoid_block":"510131034011016","dist_mi":0.9471},
<pre>{'geoid_bg': 510131016033', geoid_block: '5101396100103', dist_mi:1.2018}</pre>	<pre>,('geoid_pg': '510131016033','geoid_block': 510139601001023','atstml'11.031,'geoid_pg': '510131016033','geoid_block': 5101310501002', 'atst_ml'10.9402',</pre>
{"geoid_bg': '510131016033', geoid_block': '510131016032004', 'dist_mi:0.5769}	,("geoid_pg': "510131016033', "geoid_block': 51013101500101", "dist_ml'1.0512),("geoid_bg': "510131016033'', "geoid_block': 51013101501002', "dist_ml': 1.02?,
{"geoid_bg': '510131016033', 'geoid_block': '510131016021001', 'dist_mi':1.0225}	,("geoid_bg': "510131016033', "geoid_block': "510131016021000'', "dist_ml': 0.9078),("geoid_bg': "510131016033'', "geoid_block'': 510131017014001', "dist_ml': 0.7449},
{"geoid_bg": "510131016033","geoid_block": "510139801001005","dist_mi":0.6934}	,("geoid_bg":"510131016033","geoid_block":"51013101500105","dist_m1":0.9815,,("geoid_bg":"510131016033","geoid_block":"510131016013","dist_m1":0.9792},
{"geoid_bg": "510131016033","geoid_block": "510139801001014","dist_mi":0.9696}	,("geoid_bg":"510131016033","geoid_block":"51013101601109","dist_m1":0.841},("geoid_bg":"510131016033","geoid_block":"510131017022101","dist_m1":0.4049),
{"geoid_bg": "510131016033","geoid_block": "510131016012009","dist_mi":0.2582}	,("geoid_bg":"510131016033","geoid_block":"510131016033000","dist_m1":0.1183),("geoid_bg":"510131016033","geoid_block":"510131017022101","dist_m1":0.435},
{"geoid_bg": "510131016033", "geoid_block": "510131017011007", "dist_mi":1.2307)	,("geoid_bg':"510131016033","geoid_block":"510131017022005","dist_ml":0.2976},("geoid_bg':"510131016033","geoid_block":"510131017022009","dist_ml":0.2976},
{"geoid_bg": "510131016033", "geoid_block": "51013801001142", "dist_mi":0.4056}	,("geoid_bg':"510131016033","geoid_block":"510131016033008","dist_ml":0.6625},("geoid_bg':"510131016033","geoid_block":"510131017022000","dist_ml":0.6669},
{"geoid_bg": "510131016033", "geoid_block": "510131017022010", "dist_mi":0.4691}	,("geoid_bg':"510131016033","geoid_block":"510131017032004","dist_ml":0.6625),("geoid_bg':"510131016033","geoid_block":"510131017022000","dist_ml":0.3097},
{"geoid_bg":"510131016033","geoid_block":"510139801001025","dist_mi:0.4264}	,("geoid_bg":"510131016033","geoid_block":"510131017032001","dist_mi":0.6906},("geoid_bg":"510131016033","geoid_block":"510139801001036","dist_mi":1.1263},
{"geoid_bg":"510131016033","geoid_block":"510131017031001","dist_mi":0.7154}	{"geoid_bg":"510131016033","geoid_block":"510131017032002","dist_mi":0.6004},("geoid_bg":"510131016033","geoid_block":"51013101601102","dist_mi":0.6408},
{"geoid_bg":"510131016033","geoid_block":"510139801001008","dist_mi":0.2021,	("geoid_bg":"510131016033","geoid_block":"510131016011000","dist_mi":0.5972},{"geoid_bg":"510131016033","geoid_block":"51013101602102","dist_mi":0.6408},
<pre>{"geoid_bg": "510131016033", "geoid_block": "510131015002000", "dist_mi":1.1167)</pre>	,("geoid_bg":"510131016033","geoid_block":"510131016032000","dist_mi":0.4516),("geoid_bg":"510131016033","geoid_block":"510131016011014","dist_mi":1.1953),
{"geoid_bg": "510131016033", "geoid_block": "510131016011017", "dist_mi":1.2349}	,{"geoid_bg":"510131016033","geoid_block":"510131004003008","dist_mi":1.0548),("geoid_bg":"510131016033","geoid_block":"510131016011002","dist_mi":0.7354},
{"geoid_bg": "510131016033", "geoid_block": "510131016011001", "dist_mi":0.9529}	,{"geoid_bg":"510131016033","geoid_block":"510131016012010","dist_mi":0.4199),("geoid_bg":"510131016033","geoid_block":"51013101601201","dist_mi":0.4557),
{"geoid_bg":"510131016033","geoid_block":"510131016012000","dist_mi:0.3739}	,("geoid_bg":"510131016033","geoid_block":"510131017023004","dist_mi":0.6506,,"geoid_bg":"510131016033","geoid_block":"510131016032001","dist_mi":0.4265},
{"geoid_bg":"510131016033","geoid_block":"510131016012003","dist_mi:0.605},	("geoid_bg":"510131016033","geoid_block":"510131016011003","dist_mi":0.417),("geoid_bg":"510131016033","geoid_block":"51013101601200","dist_mi":0.7583},
{"geoid_bg":"510131016033","geoid_block":"510131017031000","dist_mi:0.666},	("geoid_bg":"510131016033","geoid_block":"510131034011014","dist_mi":1.1309},{"geoid_bg":"510131016033","geoid_block":"510131017022003","dist_mi":0.4243},
<pre>{"geoid_bg":"510131016033","geoid_block":"510139801001002","dist_mi:0.4221}</pre>	,("geoid_bg":"510131016033","geoid_block":"510139801001024","dist_mi":0.6265),("geoid_bg":"510131016033","geoid_block":"510131015001004","dist_mi":1.2164),
{"geoid_bg":"510131016033","geoid_block":"510131018021002","dist_mi":1.0928}	,{"geoid_bg":"510131016033","geoid_block":"510131017011003","dist_mi":1.1649),("geoid_bg":"510131016033","geoid_block":"510131015002013","dist_mi":1.2489),
{"geoid_bg":"510131016033","geoid_block":"510131016011006","dist_mi":0.944},	("geoid_bg":"510131016033","geoid_block":"510131016011015","dist_mi":1.0032),{"geoid_bg":"510131016033","geoid_block":"510131016011007","dist_mi":1.2489),
<pre>{"geoid_bg":"510131016033","geoid_block":"510131016011018","dist_mi:0.942),</pre>	{'geoid_bg":"510131016033","geoid_block":"510131016011012","dist_mi":0.7707},{'geoid_bg":"510131016033","geoid_block":"510131016011011","dist_mi":0.7364},
{"geoid_bg":"510131016033","geoid_block":"510131016022000","dist_mi:0.783},	{'geoid_bg":"510131016033","geoid_block":"510131017013000","dist_mi":0.9547},{'geoid_bg":"510131016033","geoid_block":"51013101601139","dist_mi":0.6522},
{"geoid_bg":"510131016033","geoid_block":"510139801001001","dist_mi":0.228};	,{'geoid_bg":"510131016033","geoid_block":"51013101501003","dist_mi":1.1531},("geoid_bg":"510131016013","geoid_block":"51013101601101","dist_mi":0.7327},
{"geoid_bg":"510131016033","geoid_block":"510131016011008","dist_mi":0.8148} {"geoid_bg":"510131016033","geoid_block":"510139801001021","dist_mi":0.9479} {"geoid_bg":"510131016033","geoid_block":"510131034011005","dist_mi":1.0321}	,("geoid_bg":"510131016033","geoid_block":"510131015001006","dist_mi":0.138},("geoid_bg":"510131016033","geoid_block":"510131016033001","dist_mi":0.138},("geoid_bg":"510131016033","geoid_block":"510131016033001","dist_mi":0.1502}, ,("geoid_bg":"510131016033","geoid_block":"5101330901001144","dist_mi":0.1708),("geoid_bg":"510131016033","geoid_block":"510131017021002","dist_mi":0.1508},
{"geoid_bg":"510131016033","geoid_block":"510131017022004","dist_mi:0.3641}	,("geoid_bg":"510131016033","geoid_block":"510131017011001","dist_mi":0.3);{"geoid_bg":"510131016033","geoid_block":"510133016032003","dist_mi":0.5898},
{"geoid_bg":"510131016033","geoid_block":"510139801001012","dist_mi:0.6886}	,("geoid_bg":"510131016033","geoid_block":"510131016032002","dist_mi":0.3};"geoid_bg":"510131016033","geoid_block":"510131017023003","dist_mi":0.5898},
{"geoid_bg":"510131016033","geoid_block":"510131017023005","dist_mi:0.7139}	,("geoid_bg":"510131016033","geoid_block":"510131011002","dist_mi":1.0057),("geoid_bg":"510131016033","geoid_block":"510131017023000","dist_mi":0.5855},"dist_mi":0.5898}
{"geoid_bg":"510131016033","geoid_block":"510131016031001","dist_mi:0.6827}	,("geoid_bg":"510131016033","geoid_block":"51013101601104","dist_mi":0.7705),("geoid_bg":"510131016033","geoid_block":"510131016031000","dist_mi":0.5841},
{"geoid_bg":"510131016033","geoid_block":"510131034011001","dist_mi:0.6897}	,("geoid_bg":"510131016033","geoid_block":"510131018021009","dist_mi":1.1356},("geoid_bg":"510131016033","geoid_block":"5101310160310014","dist_mi":1.04},
{"geoid_bg":"510131016033","geoid_block":"510131018023000","dist_mi":1.2418}	,("geoid_bg":"510131016033","geoid_block":"510131016021001","dist_mi":0.9734},("geoid_bg":"510131016033","geoid_block":"510131017013002","dist_mi":1.071},
{"geoid_bg":"510131016033","geoid_block":"510131018021001","dist_mi":1.0394}	,("geoid_bg":"510131016033","geoid_block":"510131018021000","dist_mi"10.9247},("geoid_bg":"510131016033","geoid_block":"510131017013001","dist_mi"1.0145},
{"geoid_bg":"510131016033","geoid_block":"510131017014002","dist_mi":0.9059	,("geoid_bg":"510131016033","geoid_block":"510131017012000","dist_mi"10.9447},("geoid_bg":"510131016033","geoid_block":"510131017014003","dist_mi"10.94145,
{"geoid_bg":"510131016033","geoid_block":"510131017011005","dist_mi":1.1335}	,("geoid_bg":"510131016033","geoid_block":"510131017011004","dist_mi"10.9447,("geoid_bg":"510131016033","geoid_block":"510131016012014","dist_mi"10.9427,
<pre>{'geoid_bg':510131016033','geoid_block':510139801001015','dist_mi:0.9252}</pre>	,("geoid_bg":"510131016033","geoid_block":"510139801001004","dist_mi":0.5766),("geoid_bg":"510131016033","geoid_block":"510139801001000","dist_mi":0.2724},
{'geoid_bg':510131016033','geoid_block':51013980100103','dist_mi':0.9737}	,("geoid_bg":"510131016033","geoid_block":"51013101601113","dist_mi":0.874},("geoid_bg":"510131016033","geoid_block":"510131014001","dist_mi":0.83466),
{'geoid_bg':510131016033','geoid_block':510131017022001','dist_mi':0.2011}	,("geoid_bg":"510131016033","geoid_block":"510131017011006","dist_mi":1.1643},("geoid_bg":"510131016033","geoid_block":"51013104011000","dist_mi":1.1538},
<pre>['geoid_bg':'510131016033','geoid_block':'510131018021010','dist_mi:1.1347)</pre>	,("geoid_bg":"510131016033","geoid_block":"510131018022000","dist_m1":1:2211,("geoid_bg":"510131016033","geoid_block":"510131017032000","dist_m1":0.4199},
{"geoid_bg':'510131016033','geoid_block':'510131017022012','dist_mi':0.7423}	,("geoid_bg":"510131016033","geoid_block":"510131017022008","dist_m1":0:2346),("geoid_bg":"510131016033","geoid_block":"510131011001","dist_m1":0.8391},
{"geoid_bg':'510131016033','geoid_block':'510131017022002','dist_mi':0.2436}	,("geoid_bg":"510131016033","geoid_block":"510131016033007","dist_m1":0.0273},("geoid_bg":"510131016033","geoid_block":"510131017014004","dist_m1":0.7692},
{ geoid_bg': 510131016033, geoid_block : 51013101/023006, ats_m::0.858}}	<pre>,('geoia_bg:'510131016033','geoia_block':51013101/02/013','atstm1:0.9451,\'geoia_bg':51013101033','geoia_block':51013101/02/00','atstm1:0.4022,</pre>
{ geoid_bg': 510131016033, geoid_block : 51013101702006, ats_m::0.862}	,("geoia_bg':"510131016033',"geoia_block':51013101703104","atstm1:0.94517,\("geoia_bg':510131016033","geoia_block':51013101104104","atstm1:0.1011,
{ geoid_bg': 510131016033, geoid_block : 510131017031003, dist_m::0.8718}	,("geoia_bg':510131016033","geoia_block':510131016012012","atstm1:0.2459},("geoia_bg':510131016033","geoia_block':510131017031002","atstm1:0.8541},
{ geola_bg : 510131016033, geola_block : 51013103401100/, dist_m::0.1263 { "geola_bg : 510131016033, "geola_block : 510131016033044", "dist_mi:0.1806 { "geola_bg : 510131016033, "geola_block : 510131016033043", "dist_mi:0.1806 { "geola_bg : 510131016033, "geola_block : 510131016033043", "dist_mi:0.1806 { modela_bg : 510131016033, "geola_block : 510131016033043, "dist_mi:0.1806 { modela_bg : 510131016033, "geola_block : 510131016033043, "dist_mi:0.1806 { modela_bg : 510131016033, "geola_block : 510131016033043, "dist_mi:0.1806 { modela_bg : 510131016033, "geola_block : 510131016033044, "dist_mi:0.1806 { modela_bg : 510131016033, "dist_mi:0.1806 { modela_bg : 510131016033, "geola_block : 510131016033043, "dist_mi:0.1806 { modela_bg : 510131016033, "geola_block : 510131016033044, "dist_mi:0.1806 { modela_bg : 510131016033, "geola_block : 510131016033044, "dist_mi:0.1806 { modela_bg : 510131016033, "geola_block : 510131016033044, "dist_mi:0.1806 { modela_bg : 510131016034, "dist_mi:0.1806 { modela_bg : 5101310004, "dist_mi:0.1806 { modela_bg : 510131004, "dist_mi:0.	<pre>,('geoia_bg:'510131016033','geoia_block':51013101601306'','atstml'11.2005/,('geoia_bg':510131016033'','geoia_block':51013101603200'','atstml'10.276), ,('geoid_bg':'510131016033'','geoid_block':51013101603200'','atstml'10.1943),('geoid_bg':'510131016033'','geoid_block':51013101603200'','atstml'10.276), ,('geoid_bg':'510131016033'', 'geoid_block':'51013101603300'','atstml'10.1276), ,('geoid_bg':'510131016033'', 'geoid_block'':51013101603300'','atstml'10.1276), ,('geoid_bg':'510131016033'', 'geoid_block'':510131016033'','atstml'10.1276), ,('geoid_bg':'510131016033'', 'geoid_bg':'510131016033'','atstml'10.1276), ,('geoid_bg':'510131016033'', 'geoid_bg':'510131016033'','atstml'10.1276), ,('geoid_bg':'510131016033'', 'geoid_bg':'510131016033'','atstml'10.1276), ,('geoid_bg':'510131016033'', 'geoid_bg':'510131016033'','atstml'10.1276), ,('geoid_bg':'510131016033'','atstml'10.1276), ,('geoid_bg':'5101310603'','a</pre>
<pre>[ geoid_pg : 51031016033 , geoid_block : 51015101/022000 , d18t_m1 :0.1009)</pre>	<pre>, 'geoid_bg': "Stol3106033", 'geoid_block': 'Stol3106012002", 'dist_mi:0.514}, ('geoid_bg': 'Stol31016033', 'geoid_block': 'Stol31016012017', 'dist_mi:0.522},</pre>
{ "geoid_pg : 510131016033 , "geoid_block": 51013101612011 , "d1st_m1 :0.2609)	,('geoid_bg': "Stol3106033", 'geoid_block': 'Stol31016012005", 'dist_mi:0.514}, ('geoid_bg': 'Stol31016033', 'geoid_block': 'Stol31010011017', 'dist_mi:0.522},
{ "geoid_pg : 510131016033 , "geoid_block": 51013101612014 , "d1st_m1 :0.4998)	,('geoid_bg': "Stol31016033", 'geoid_block': 'Stol31016012005", 'dist_mi:0.3371}, ('geoid_bg': 'Stol31016033', 'geoid_block': 'Stol31016012017', 'dist_mi:0.522},
{ "geoid_pg : 510131016032 , "geoid_block": 510130161010101 , "d1st_m1 :0.4998]	)
<pre>geta_bg : station (0.3206) {"geoid bg : "\$10131016033", "geoid block": "\$1013101723002", "dist_mi":0.5253 {"geoid bg": "\$10131016033", "geoid block": "\$1013101632003", "dist_mi":0.5253 {"geoid bg": "\$10131016033", "geoid block": "\$1013101632003", "dist_mi":0.31911 "geoid bg": "\$10131016033", "geoid block": "\$1013101632006", "dist_mi":0.3016 "geoid bg": "\$1013101603", "geoid block": "\$1013101632006", "dist_mi":0.3016 "geoid bg": "\$1013101603", "geoid block": "\$1013101632006", "dist_mi":0.3016 "geoid bg": "\$1013101603", "geoid block": "\$1013101632006", "dist_mi":0.3016 "geoid bg": "\$101310000", "dist_mi":0.3016 "geoid bg": "geoid bg"; "geoid bg"; "geoid bg"; "g</pre>	<pre>, 'geoid_bg': "Stol31016033", 'geoid_block': 'Stol3101012008', dist_mi '10.5872), ('geoid_bg': "Stol31016033", 'geoid_block': 'Stol310160330", 'dist_mi '10.5872), ('geoid_bg': 'Stol31016033", 'geoid_block': 'Stol310160330", 'dist_mi '10.5872), ('geoid_bg': 'Stol31016033", 'geoid_block': 'Stol310160330", 'dist_mi '10.5872), ('geoid_block': 'Stol310160330", 'dist_mi '10.5874), ''''''''''''''''''''''''''''''''''''</pre>
{ geoid_bg : 510151016033", geoid_block : 510131016032006", "dist_m1":0.3891}	, (geota_pg : 510151018033 , geota_block : 510151018021011 , alst_ml :1.144},{ geota_pg : 510151018033 , geota_block : 510151015002001 , dist_ml :1.1904},
{"geoid_bg :: 510131016033", "geoid_block :: 510131015002014", "dist_m1":1.1564}	]



### Data Visualization & Sharing Project Wikis (XWiki)

SOCIAL & DECIS Biocomplexity Institute, Univ	VION ANALYTICS DIVISION	Q ♣ 📃 ≡
Applications	🕷 * / Projects - Active * / / Meetings * / SDAD Planning and Meetings * / 2018-11-7 meeting *	Tips
<ul> <li>み Blog</li> <li>☆ Calendar</li> <li>≕ Dashboard</li> <li>❹ FAQ</li> </ul>	2018-11-7 meeting Last modified by Aaron Schroeder on 2019/01/31 18:02	Jump to any page in the wiki with Ctrl+G or Meta+G. See more shortcutst?.
<ul> <li>File Manager</li> <li>Forums</li> <li>Help</li> <li>Ideas</li> </ul>	USDA ERS Broadband meeting November 7, 2018 at 11 am	My Recent Modifications
<ul> <li>Meetings</li> <li>i Menu</li> <li>I Polls</li> <li>■ Sandhoy</li> </ul>	Team – Sallie, Stephanie, Aaron, Josh, Teja, Devika Discuss SOW - Sallie Objective Discuss	Population Demographic Counts Household & Rental Counts Education Index
<ul> <li>Task Manager</li> <li>More applications</li> </ul>	Deliverables     Roles Topics to research – learn about and summarize for others on the team to also get smart	Soil Productivity Index Land Cover Acreage
Navigation	<ul> <li>RUS program and data sources - https://www.rd.usda.gov/about-rd/agencies/ruia-utilines-service/2</li> <li>John Pender also mentioned this program as providing insights about RUS (Rural Business Cooperative Service criteria. https://www.rd.usda.gov/about-rd/agencies/rural-business-cooperative-service/2</li> <li>They might be able to get some feedback on how they see the criteria, how they rank projects, etc.).</li> </ul>	Need help?
> Blog	<ul> <li>John Pender is attempting to get Data Dictionary for RUS data</li> <li>John can share some recent reports/articles on this program. See email below.</li> </ul>	you can contact:
• FAQ • Forums > Help • Home	<ul> <li>Deep dive into 2 to 3 geographic areas – at Nov. 7 meeting, brainstorm on criteria for selection What areas in Virginia might be candidates? Identify X areas and provide overview of the area and rank areas for selection. "Select a few geographic regions initially that have rich data sources and then progress."</li> <li>Initial under of the Reserver</li></ul>	Community Support     Professional Support
Menu     Projects - Active     Sub-County Data for	Minual review of the increative – start with Elefature Folder     Are three studies that have examined broadband deployment and impact on property values?     What is current state of broadband diffusion across the US?     What are characteristics of areas that do not have BB?	
Policy ~ USDA ERS Broadband ~ Meetings ~ SDAD Planning and	<ul> <li>winat data sources are identified in the ancies?</li> <li>Other questions we want to know from literature?</li> <li>Describe Data Sources - keep in mind, we are looking at housing AND business property values - develop a data model (see examples from Roanoke and USDA AFRI proposal on Rural Entrepreneurship -both are in Related Projects folder)</li> <li>FCC broadband data - learn about different sources (they vary over time)</li> </ul>	
Meetings • 2018-11-27 SDAD planning • 2018-11-7 meeting • 2018-11-7 meeting 2 LISDA Meetings	CoreLogic - USDA has CL data from 2005 to 2015 for metro and non-metro Areas.     National Employment Time Series Database - https://www.minneapolisfed.org/institute/working-papers/an-assessment-of-the-national- establishment-time-series-database c     Dun and Bradstreet data?     American Community Survey data – what data would be relevant to this study?     Other data relevant to the project?	
> Presentations	Action items.	
• xi	Read materials in MUST READ folder. Add materials to this folder that ALL team members should read.	



### Data Visualization & Sharing Project Wikis (XWiki)



UNIVERSITY

### Thank You!

