🏠 **techred home** > data anlaytics master sequence

Three-course data analytics series at CCAC's North Campus

1. DAT-102: Introduction to Data Analytics

2. DAT-201: Data Analytics 1 [Taught by Coral Sheldon-Hess only SP20]

Course concept progression

The following table maps course session dates, lesson topics, references, and content links for all three Data Analytics courses in the series.

course	date	wk no.	session links	learning objectives	out-of-class work
DAT-102	Tue 1-SEP-2020	1	**Introduction to data analytics** Recording of zoom session: Data analytics progression	TR.102.DS.3.A - Decompose the data analytics field TR.102.DS.1.A - Data Tables - Creating: **Create a data table with logically assigned types for each column and a unique identifier for each row**	Please develop a "strip survey" containing a categorical question and a opinion/spectrum question. Compose the tiny survey in a text document and upload to a folder named with your public ID in our shared drive. ✏️ PDF summary of strip surveys from FA18 Navigate to Strip surveys then Fall 2020 ☁️ DAT-102 Project upload shared drive⤴

DAT-102: FA20: Session 1: Course Basics

▶️

course	date	wk no.	session links	learning objectives	out-of-class work
			Types of data and strip survey DAT102-FA20... 📖 Course syllabus 🖨 Week 1 Station Guide 🖨 CCAC Data Analyics AS Overview 🖨 CCAC Data Analytics Cert Overview		
DAT-102	Tue 8-SEP-2020	2	**Session recordings (FA20)** Part 1: Data encoding DAT102-Session2-DataEncoding No views • Sep 10, 2020 Part 2: Data structures overview DAT102-Sess2-DataStructuresOverview No views • Sep 10, 2020	Broadly Classify data analytic artifacts/products /displays (Quant/qual/categorical/textual) TR.102.DS.3.C - Continuous & categorical variables TR.102.DS.3.D - Data structures (list, set, stream, table, graph, tree) TR.102.DS.3.E - Analytic modes: describing, modeling, predicting TR.102.DS.1.B - Data Tables - Converting: **Export and import data tables in .xslx, .ods, .csv formats**	*Fall 2020* 1. Finish your graph, upload to onedrive 2. Choose a graph that's interesting to you, and create a tabular representation, either on paper or in a spreadsheet. Try to encode as much of the original data as you can (i.e. do the edges have additional meaning beyond just "I'm an edge", do the nodes have values? Do they have types) 3. Save your tabular representation using only your first name, not the name of the creator. Save it in the special directory called "Fall 2020_tables_ANONYMIZED_notopicorcrea 4. If you didn't make a strip survey, finish

course	date	wk no.	session links	learning objectives	out-of-class work
			Data structures stations Binary-Hex-Dec Converter *Graph exercise:* With all your representations complete, open our shared upload directory below. Create a new directory named with your first name and the topic of your data. Upload an image file of your graph into the directory. Graph Upload OneDrive Location encoding gsheet Gephi download Gephi quick start		that and upload in the link above this cell. Data structures home practice (pre-COVID)
DAT-102	Tue 15-SEP-2020	3	Ida Mae Darsow Interest Inventory Results Photos of Ida Mae Non-summary statistics Strip survey cloud drive [Pre-COVID Quant variable profile Editable Quant variable profile PDF Online box plot image creator Sample strip survey analysis		1. Create your strip survey master drawing in the shared google drive 2. By Friday 18-Sep @ midnight please have submitted responses for each of your peer's strip surveys in their respective directories. 3. Starting Sat morning, and before class starts next week, please create a spreadsheet in your strip survey folder on google drive, with each survey response getting its own row/record in the table. Give each survey a unique identification number, which you can use to check your data in the

course	date	wk no.	session links	learning objectives	out-of-class work
					spreadsheet.
DAT-102	Tue 22-SEP-2020	4	**Strip survey analysis** 🚗 Strip survey results Shared GDOC ▦ Sample analysis sheet from 22-SEP ☁ Strip survey cloud drive (Deprecated) 📊 Box plot generator **Summary-based descriptive stats: mean and standard deviation** 🖨 Distributions and variance (under const) 🖨 Quant profile V.1.0 *Extra* 🖨 Distributions playground spreadsheet 🖨 Edgewood and swissvale comparison 🕐 **instructor post-session notes**	1. Record student responses to your strip survey in a google sheet inside your google drive directory 2. Measure your total line length. Enter this value in a dedicated special cell in your spreadsheet to use for scaling. 3. Compute a scaled score for your slicer in the spreadsheet as a Percent of total line length. Do this by adding a new column to the right of your raw measured value. 4. Use formula master skills to generate a percent of total line distance. Don't forget an absolute reference to your total line length 5. With scaled values, compute your quant profile for your aggregate responses (not sliced) 6. Create new tabs in your spreadsheet, one for each of your possible slicer responses. name the tabs logically, without spaces or weird characters 7. Copy your aggregate data from your first sheet into each of your slicer tabs 8. Select all your data and sort the data by slicer question response. Delete the rows of the responses whose slicer answer is NOT the focus of that tab 9. With your responses trimmed by slicer, compute your variable profile values for each of your data sub-sets (N, min, median, max, lower fence, upper fence, left whisker, right whisker) 10. With those compute values in place, use our unified box and whisker tool to create box plots for your aggregate and sliced responses 11. Right click the resulting image in the box plot tool and save them to your local drive. Then upload them with sensible names to your google drive strip survey directory 12. We'll do the group analysis next week.	

course	date	wk no.	session links	learning objectives	out-of-class work
DAT-102	Tue 29-SEP-2020	5	Phase 1: Spreadsheet play-along: center and spread computation and manipulation Phase 3: Trade-offs and conflicting priorities group exercise Phase 4: Debrief and discussion of normality assumptions in statistical inference		Complete activities 1A - 1K in Chapter 1 of Statistics Notes handout The key for the exercises will be posted here during class next week. 🖨 Summarizing Data: Ch 1: KEY (PDF)
DAT-102	Tue 6-OCT-2020	6	**Applying mean, median, and standard deviation** Match up the Distribution, stats blocks, box plot, and data source in this file Phase 1: Reviewing key concepts from stats packet & real-time data gathering and analysis Phase 3: Group and dispute exercise: connecting distributions, summary stats, and data-backed claims Phase 4: Internalizing the concept of the standard normal curve	TR.102.DS.6.A - Surveys - Designing: TR.102.DS.6.B - Surveys - Sampling & Administering: TR.102.DS.6.C - Surveys - Analyzing:	
DAT-102	Tue	7	**Sampling!**		Please study the two American Journal of Public Health articles distributed in class. Prepare to dig

course	date	wk no.	session links	learning objectives	out-of-class work
	13-OCT-2020		Begin library section sampling, to be continued next week. 📑 Library sampling project		into their confience intervals for each sub-population: 1. Law Enforcement Agencies' Perceptions of the Benefits of and Barriers to Temporary Firearm Storage to Prevent Suicide (Feb-2019, Am J. Pub Health) by Brooks-Russell, Ashley; Runyan, Carol; Betz, Marian E.; Tung, Greg; Brandspigel, Sara; Novins, Douglas K. ⤢ 2. Sociodemographic Correlates of Electronic Nicotine Delivery Systems (ENDS) Use in the US (Sep-2019, Am J. Pub Health), by Spears, Claire Adams; Jones, Dina M.; Weaver, Scott R.; Huang, Jidong; Yang, Bo; Pechacek, Terry F.; Eriksen, Michael P. (2016-2017) ⤢
DAT-102	Tue 20-OCT-2020	8	Built-in flex time		

course	date	wk no.	session links	learning objectives	out-of-class work
DAT-102	Tue 27-OCT-2020	9	**Library samples continued** 🔖 Library sampling project 🔲 Class project tracker Shared Google Drive ⤢ 🔒 StatKey online data cruncher ⤢	Sampling 1: Implement the process of making an inference about a population parameter from a sample. Sampling 2: Use a statistical package--such as StatKey--to experimentally estimate the standard error of the sampling distribution	**NOTE: Skip hypothesis testing questions/sections** Dedicate a few hours hours to carefully responding to the analysis questions from your library sample. See our sampling module, and choose the library sampling mini-project. Uplod all your work in our Shared drive for library upload also linked in the module resources. Be sure to generate your own file prefix to ensure grouping of your work when the directory is sorted.
DAT-102	Tue 3-NOV-2020	10	**Interpreting sample data** 🔖 Library sampling project 🔲 Class project tracker GSheet ⤢ *Session agenda* 1. Sampling real-time socrative exercise (rm. name = DARSOW) 2. Two parameter types: mean and proportion 3. Mystery population exercise 4. Preview of out-of-class work: Opportunity Atlas investigations		**Wrap-up library sampling** Please follow the out-of-class assignment instructions from last week if you didn't yet complete a thorough working through of questions 1-6 of our analysis guide. Remember: no hypothesis tests at this stage. And then upload your work to the shared drive linked in last week's HW.

course	date	wk no.	session links	learning objectives	out-of-class work
DAT-102	Tue 10-NOV-2020	11	🎬 Screen cast of class session CI review; opp atals overview **Opportunity Atlas mini-project: multi-type data policy inquiry** 🔲 Exploring the Opportunity Atlas 🖨 American factfinder home 🖨 American Community Survey Error Rates Explained	TR.102.DS.7.A - Experiments - Designing: TR.102.DS.7.B - Experiments - Treatment assignment & Implementing: TR.102.DS.7.C - Experiments - Analyzing: TR.102.Q.10 - Standard errors TR.102.Q.11 - Student's T-tests - Setup TR.102.Q.12 - Student's T-tests - Interpretation	**Dig into the Opp Atlas** Please complete the exercises 1 and 2 on the Exploring the Opportunity Atlas and upload your results to our shared drive when complete. Be sure to print off the student worksheet (or edit it digitally) linked inside the module. *Est. Time: 3-ish hours* The true/false exercise in the student worksheet is very rigorous and worthy of some thought. Dedicating beyond 3 ish hours to this assignment is not intended, so please do not stress about "not finishing". I'd rather you take your time and explore the Atlas than worry about the status of your answers to questions on a worksheet. In other words, the worksheet is our means of familiarity and not meant to be an assignment in its own right. *Start thinking about your final project* 📊 DAT-102 Final project specs
DAT-102	Tue 17-NOV-2020	12	**Opp Atlas 2** 👍👎 Data-based program evaluation **Final project practice and design** 📊 DAT-102 Final project specs	[1]	**OPTIONAL Out of class:** *Digest PGH Inquality report* Due to COVID-19 reorganiation, we will be unable to discuss the data and the sociology behind Pittsburgh's Inequality Across Gender and Race Report issued by the Pittsburgh Gender Equity Commission. As you desire, please engage with the report on your own and with others in your various circles. These discussion questions may be a guide for your discussion: 1. Review the study's aggregation of

course	date	wk no.	session links	learning objectives	out-of-class work
					smaller racial subcategories into the "AMLON" category. What are the advantages of this statistical approach? Its limitations? Would there be other ways to aggregation races into smaller categries? 2. Review the Report's focus areas in the section called "Cultivating Livability." Which of these priorities do you believe are most salient at this time in Pittsburgh? Most data-based? Least data-based? 3. Carefully study the comparison methodology in Appendix A. Develop a thoughtful opinion of the author's assertion on page 72, third paragrah which starts: "When outcomes, like grade reten tion rates, are similar across cities they are likely to be driven more by national policies and factors...". Can you think of any indicator patterns which do not exhibit this behavior?
DAT-102	Tue 24-NOV-2020	-	TURKEY DAY BREAK ALL WEEK		
DAT-102	Tue 1-DEC-2020	13	❓ Experimental design Final project concept development	1	1

course	date	wk no.	session links	learning objectives	out-of-class work
DAT-102	Tue 8-DEC-2020	14	**FINAL EXAM PERIOD from 6:00 - 8:00 pm** 📊 Final session checklist (editable) and or PDF 📊 DAT-102 Final project Board layout 📊 DAT-102 Final project Group Tracker 📊 DAT-102 Final project specs 📊 Final project cloud upload		

Data 201: Data Analytics 1

Not offered by Eric Darsow in Spring of 2020 (rather by Professor Coral Sheldon-Hess)

course	date	wk no.	session links	learning objectives	out-of-class work
DAT-201	TUE 03-SEP-19	1	**Session outline:** 1. Welcome and introductions 2. Project-based learning in action: Review of past term projects: project repository and student response sheet 3. Syllabus review 4. Pivot table glory: Past example 5. Pivot table glory: Your	SPDSHT1: **Implement VLOOKUP formulas in spreadsheets** SPDSHT2: **Fomulate a spreadsheet to properly get slurped up by a pivot table** SPDSHT3: **Create a pivot table to answer inquiry questions by configuring row and column selections**	

course	date	wk no.	session links	learning objectives	out-of-class work
			turn! Grade comparison. 🖨 Session guide:VLookup() and Pivot tables review ✅ Shared drive of grade analyses		
DAT-201	TUE 10-SEP-19	2	**Map projections and Intro to QGIS** 🖨 Session guide: Maps 1	TR.201.DS.8.A - Maps - Projections TR.201.DS.8.B - Maps - Vector (points, lines, and polys) & raster (bands) TR.201.DS.8.C - Maps - QGIS fundamentals	**Part 1: Pre-reading for week 2: Maps!** Pre-reading on Responsible map making **Part 2: Install QGIS** QGIS install homepage by platform. This software package is large and complicated, but has been ported to Windows and OSX. Many students have no problems with the install, but in some cases, there are dependency issues that take quite a bit of time to resolve because QGIS is based on python and several other packages. Please follow the instructions carefully and have a working copy on your computer by 10-SEP-19 for in-class demo (but realistically, the 17th is when we'll start using it in class). **Homework:** Explore QGIS, make sure you understand what a layer is and how to add one. Come with questions next week. For anyone who doesn't want to aimlessly explore, here's a good (but fast!) video introduction to QGIS.

course	date	wk no.	session links	learning objectives	out-of-class work
DAT-201	TUE 17-SEP-19	3	**QGIS Demonstrations** 🖨 Session guide: Maps 2	TR.201.DS.8.D - Maps - Creating study areas TR.201.DS.8.E - Maps - Flat Joins TR.201.DS.8.F - Maps - Spatial Joins	**Homework:** Details available on the session guide; short version: make a map with PASDA data (mostly in-class), and start on your mid-semester mapping project (mostly out-of-class). Be ready to share what you're planning to do and any initial steps you've taken, next week.
DAT-201	TUE 24-SEP-19	4	**Mapping with Nine Mile Run Watershed Association** 🔗 NMRWA (nonprofit org; homepage) Solve real-world problems with a local nonprofit! 📊 Watershed dataset (both zip & uncompressed)		
DAT-201	TUE 01-OCT-19	5	**QGIS and Map Layouts** 🖨 Session guide: Maps 3	TR.201.DS.8.G - Maps - Layouts & printing TR.201.DS.8.H - Maps - Web compatability	• Download Open Refine, and make sure it's up and running on your machine. • Get your mapping project started (we'll make some time for project troubleshooting in class next week). • Watch these three videos (1, 2, 3) and start playing with Open Refine.
DAT-201	TUE 08-OCT-19	6	**Work time on projects and open refine** Tutorial set of nuclear explosions dataset Student practice nuclear	CLI.FUND.1 Differentiate between the unix BASH, Microsoft Corporation's command prompt, and the Apple terminal in terms of origins, function, use, and proprietary status CLI.FUND.2 Navigate a diredctory structure with cd, ls, tab completions, and the use of the files named . and ..	1

course	date	wk no.	session links	learning objectives	out-of-class work
			explosions dataset Open refine documenation	CLI.FUND.3 Maniplate files and directories safetly with mkdir, mv, rm, and cp CLI.FUND.4 Parse file access permissions info as displayed by ls -al and safely issue commands with superuser powers via sudo	
DAT-201	TUE 15-OCT-19	7	Worktime and presenting mapping mini-project 6-7pm: Finalize mapping mini-project 7-?pm: Present project to class with feedback 🗄 **session notes**	TR.201.DS.9.E - Clients - Feedback presentations	1
DAT-201	TUE 22-OCT-19	8	Database configuration ▦ PostgreSQL module	TR.201.DB.1: Database use cases TR.201.DB.2: Types (File, relational, NOSQL) TR.201.DB.4.A - Tables - Data types TR.201.DB.4.B - Tables - Keys TR.201.DB.4.C - Tables - Foreign Keys TR.201.DB.5.A - Queries - SELECT	Unless progress in class is slower than expected, please attempt the query challenges in the last section of our postgreSQL module and be prepared to share your results with your peers next week.
DAT-201	TUE 29-OCT-19	9	Databases continued Overview of core linux tools: • getting help with man XXX • user@host notation	TR.201.DB.4.D - Tables - Manipulating TR.201.DB.6.A - Data - INSERT TR.201.DB.6.B - Data - UPDATE TR.201.DB.5.B - Queries - FROM (Joins) TR.201.DB.5.C - Queries - WHERE	Please copy in the jail census flat file, and attempt the sample quriers in our postgres guide Choose another flat file, perhaps one from the wprdc.org (hopefully, a really really big one), create a receiving table in postgres into which you copy the contents of the flat file for querying. Identify at least one compelling question you can answer using SQL statements to share with the class next

course	date	wk no.	session links	learning objectives	out-of-class work
			• port numbering • ssh tools: ssh -f for forwarding, sshfs • command line tools: head, tail, cat • remote mounting of drives	TR.201.DB.5.D - Queries - ORDER BY TR.201.DB.3: Leading vendors TR.201.DB.7 - Exporting TR.201.DB.8.A - Connecting - Spreadsheets TR.201.DB.8.B - Connecting - Python & Java	week.
DAT-201	TUE 05-NOV-19	10	Databases: Designs, features, & use cases 🔒 Postgres Modules	TR.201.DB.10.A - Design - Methodologies TR.201.DB.10.B - Design - Creating from data statements TR.201.DB.10.C - Design - Normalization TR.201.DB.10.D - Design - Many-to-many relationships TR.201.DB.10.E - Design - Spotting traps	Please devote a few hours to completing this command line exercise. you will want to secure a meaningful BASH command reference on line. Look for resources with not many ads, or ones with a .edu extension. This exercise will ask you to answer lettered questions--please record answers to them as you progress through the exercises. Also, please remember to take your time and read the man pages for commands that you aren't familiar with, such as wc and others. Command line practice Also, please start in on our postgres mini-project found with the button called "postgres mini-project" in our postgres module page.

course	date	wk no.	session links	learning objectives	out-of-class work
DAT-201	TUE 12-NOV-19	11	**PostGIS in action** 🔒 Postgres Modules See steps in "postgres mini-project outline"	TR.201.DB.9.A - Server - User configuration & permissions TR.201.DB.9.B - Server - Access, GUIs, and SSH TR.201.DB.9.D - Server - Indexes & query optimization TR.201.DB.5.E - Queries - Functions TR.201.DB.5.F - Queries - Fuzzy matching	📄 DAT-201 Final Project Guide
DAT-201	TUE 19-NOV-19	12	📄 DAT-201 Final Project Guide **Database server configuration** Carrying out even small administration tasks correctly on a database requires a basic foundation in how the larger DB system works with the operating systems and its users. **Project work time** 💾 PostgreSQL Module 1. Creating data system flow diagram & work process logs 2. Troubleshooting postgreSQL /copy commands 3. Writing queries with aggregate functions	1	1

course	date	wk no.	session links	learning objectives	out-of-class work
			and GROUP BY for analytics		
	TUE 26-NOV-19	-	THANKSGIVING BREAK!		
DAT-201	TUE 03-DEC-19	13	**MEET AT Monroeville Gov't Center 2700 Monroeville Blvd, Monroeville, PA 15146** **Tentative:** Digital meeting with Mark Egge of High Street Consulting 🧩 Jupyter notebook, PostGIS, Census exercise from Mark Egge **Collaborative project worktime & overview** Please bring questions, your data, computers, and enthusiasm for collaborative help.	TR.201.DS.9.A - Clients - Client interviews & problem scoping TR.201.DS.9.B - Clients - Specification negotiation TR.201.DS.9.C - Clients - Work process logs & billing	[1]
DAT-201	TUE 10-DEC-19	14	**Final project sharing!** 📄 DAT-201 Final Project Guide Bring fully-baked final project to class at our normal 6:00 pm. We'll share what you've discovered,	TR.201.DS.9.D - Clients - Feedback conversations TR.201.DS.9.E - Clients - Feedback presentations TR.201.DS.9.F - Clients - Tool maintenance	

course	date	wk no.	session links	learning objectives	out-of-class work
			submit grade proposals, and offer final program feedback.	planning: TR.201.DS.9.G - Clients - Iterative tool development:	