












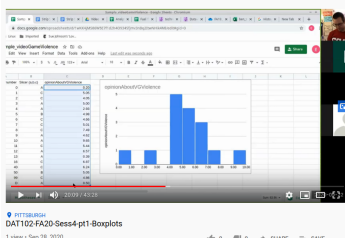
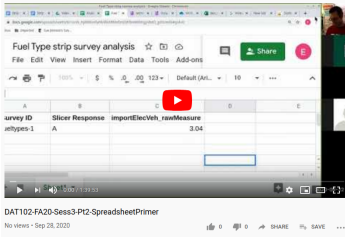
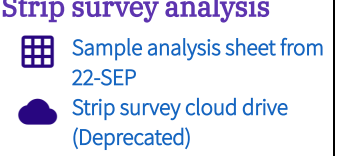

















DAT-102: Introduction to Data Analytics






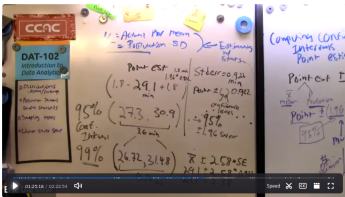


The following table maps course session dates, lesson topics, references, and content links for DAT-102, Introduction to Data Analytics

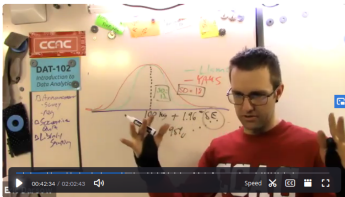







course	date	wk no.	session links	learning objectives	out-of-class work
DAT-102	Sat 6-FEB'21	1	<p>Introduction to data analytics</p> <ul style="list-style-type: none"> 🎬 SP'21 Session Recording: 6-FEB'21 ☁️ DAT-102: Course syllabus (SP'21) 😊 Name registration form 😊 Google account registration form 🖨️ CCAC Data Analytics Programm Website 📅 SP'21 Master Tracker Spreadsheet 🕒 Reaction Time Test <p>archive</p>	<p>TR.102.DS.3.A - Decompose the data analytics field</p> <p>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</p>	<p>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.</p> <p> PDF summary of strip surveys from FA18</p>
DAT-102	Sat 13-FEB'21	2	<p>Session recordings (FA20)</p> <p>Part 1: Data encoding</p>  <p>Part 2: Data structures overview</p>	<p>Broadly Classify data analytic artifacts/products/displays (Quant/qual/categorical/textual)</p> <p>TR.102.DS.3.C - Continuous & categorical variables</p> <p>TR.102.DS.3.D - Data structures (list, set, stream, table, graph, tree)</p> <p>TR.102.DS.3.E - Analytic modes: describing, modeling, predicting</p> <p>TR.102.DS.1.B - Data Tables - Converting: Export and import data tables in .xlsx, .ods, .csv formats</p>	<p><i>Fall 2020</i></p> <ol style="list-style-type: none"> 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

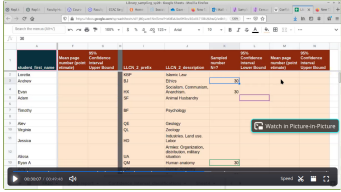
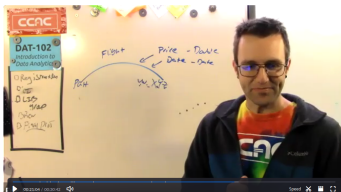
course	date	wk no.	session links	learning objectives	out-of-class work
			 <p data-bbox="636 370 982 402"> Data structures stations Binary-Hex-Dec Converter </p> <p data-bbox="636 537 982 570"><i>Graph exercise:</i></p> <p data-bbox="636 602 982 824"> With all your representations complete, open our 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. </p> <p data-bbox="636 849 982 881"> Graph Upload OneDrive </p> <hr/> <p data-bbox="636 922 982 954"> Location encoding gsheet Gephi download Gephi quick start </p> <p data-bbox="800 1057 947 1097" style="border: 1px solid black; padding: 2px; display: inline-block;">  archive </p>		<p data-bbox="1528 164 2011 256"> creator. Save it in the special directory called "Fall 2020_tables_ANONYMIZED_notopiccreator" </p> <p data-bbox="1497 280 1955 375"> 4. If you didn't make a strip survey, finish that and upload in the link above this cell. </p> <p data-bbox="1514 415 1955 448"> Data structures home practice (pre-COVID) </p>

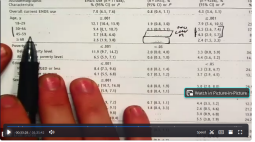


course	date	wk no.	session links	learning objectives	out-of-class work
DAT-102	Sat 20-FEB'21	3	<ul style="list-style-type: none">  Ida Mae Darsow Interest Inventory Results  Photos of Ida Mae  Non-summary statistics  Strip survey cloud drive [Pre-COVID]  Quant variable profile Editable  Quant variable profile PDF  Online box plot image creator  Sample strip survey analysis <div style="border: 1px solid black; padding: 2px; display: inline-block; margin-top: 10px;">  archive </div>		<ol style="list-style-type: none"> 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 spreadsheet.
DAT-102	Sat 27-FEB'21	4	<p><i>Session videos</i></p> <div style="display: flex; align-items: center;">    </div> <p>Strip survey analysis</p> <ul style="list-style-type: none">  Sample analysis sheet from 22-SEP  Strip survey cloud drive (Deprecated) 	<ol style="list-style-type: none"> 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) 	


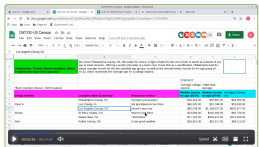









course	date	wk no.	session links	learning objectives	out-of-class work
			<p> Box plot generator</p> <p>Summary-based descriptive stats: mean and standard deviation</p> <p> Distributions and variance (under const)</p> <p> Quant profile V.1.0</p> <p><i>Extra</i></p> <p> Distributions playground spreadsheet</p> <p> Edgewood and swissvale comparison</p> <div style="border: 1px solid black; padding: 5px; display: inline-block;">  instructor post-session notes </div>	<p>10. With those compute values in place, use our unified box and whisker tool to create box plots for your aggregate and sliced responses</p> <p>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</p> <p>12. We'll do the group analysis next week.</p>	
DAT-102	Sat 6-MAR'21	5	<p><i>Session recording</i></p> <div style="text-align: center;">  </div> <p> Strip survey analysis Google Doc</p> <p> Reaction Time Test</p> <p> Reaction Time Results</p> <p> Normal distribution scales</p> <p><i>Lock^5 Book sections</i></p> <p>Chapter 2, Sections 1-Sections 4</p> <div style="border: 1px solid black; padding: 5px; display: inline-block; margin-top: 10px;">  archive </div>	<p>Draw conclusions about a data set based on box plots</p> <p>Compute the standard deviation of a data set, interpret the results, and make inferences using Z-scores</p>	<ul style="list-style-type: none"> If you didn't get a chance to finish your section of the strip survey analysis or analyze a peer's data, please do so this week. Complete activities in Chapter 1 of Statistics Notes handout <p>NOTE: Several pages are in inverted order! (9 before 8, etc.)</p> <p> Summarizing Data: Ch 1 (PDF)</p>





course	date	wk no.	session links	learning objectives	out-of-class work
DAT-102	Sat 13-MAR'21	6	<p>Applying mean, median, and standard deviation</p> <p>Match up the Distribution, stats blocks, box plot, and data source in this file</p> <p> Session Recording via Zoom</p> <p>Video Note: Password available from instructor or class peers</p> <p> Dist. Challenge Key (JPG)</p> <p> Distribution organizer (ODT)</p> <p> archive</p>	<p>TR.102.DS.6.A - Surveys - Designing:</p> <p>TR.102.DS.6.B - Surveys - Sampling & Administering:</p> <p>TR.102.DS.6.C - Surveys - Analyzing:</p>	<p> Summarizing Data: Ch 1: KEY (PDF)</p>
DAT-102	Sat 20-MAR'21	7	<p><i>Session Recording</i></p>  <p>Sampling!</p> <p>Begin library section sampling, to be continued next week.</p> <p> Library sampling project</p> <p> archive</p>		<p>Please sample 30 books from each of your two library sections: record the call number, number of pages, and some creative variable for each book in each section.</p>









course	date	wk no.	session links	learning objectives	out-of-class work
DAT-102	Sat 27-MAR'21	8	<p><i>Session recording</i></p>  <p>DAT Planning Survey  DAT Planning Survey</p> <p>Library samples continued</p> <ul style="list-style-type: none">  Library sampling project  Class project tracker Shared Google Drive   StatKey online data cruncher  <p style="text-align: center;"> archive</p>		<p>NOTE: Skip hypothesis testing questions/sections</p> <p>Dedicate a few hours to carefully responding to the analysis questions from your library sample. See our sampling module, and choose the library sampling mini-project. Upload 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.</p>
DAT-102	Sat 3-APR'21	-	Spring break; No class all week		

course	date	wk no.	session links	learning objectives	out-of-class work
DAT-102	Sat 10-APR'21	9	<p data-bbox="632 191 951 261"><i>Session Recording: Pre-groupwork</i></p>  <p data-bbox="632 521 951 591"><i>Session Recording: Post-groupwork</i></p>  <p data-bbox="632 854 951 924">Review of CI Fundamentals</p> <p data-bbox="632 951 951 1021">Review Library Sample Findings</p> <p data-bbox="632 1029 951 1057">Library sampling project</p> <p data-bbox="800 1081 951 1122">archive</p>	<p data-bbox="1001 191 1451 277">Sampling 1: Implement the process of making an inference about a population parameter from a sample.</p> <p data-bbox="1001 302 1451 388">Sampling 2: Use a statistical package--such as StatKey--to experimentally estimate the standard error of the sampling distribution</p>	<p data-bbox="1499 199 1913 227">Wrap a bow on library sampling</p> <p data-bbox="1499 245 1976 331">Complete as much as feasible of the library analysis questions and data sheets and upload them to our shared drive.</p> <p data-bbox="1499 371 1839 399">Conf. Interval article study</p> <p data-bbox="1499 417 1986 503">Please study the two American Journal of Public Health articles distributed in class. Prepare to dig into their confidence intervals for each sub-population:</p> <ol data-bbox="1499 544 1976 1024" style="list-style-type: none"> <li data-bbox="1499 544 1976 769">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. <li data-bbox="1499 797 1976 1024">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)

course	date	wk no.	session links	learning objectives	out-of-class work
DAT-102	Sat 17-APR'21	10	<p><i>Session recording</i></p>  <p>passcode: %t3chnology%</p> <p><i>Review of ENDS article confidence intervals</i></p> <p>Socrative quiz</p> <p><i>Log our final project ideas</i></p> <p> DAT-102 Final project Group Tracker</p> <p></p>		<p>Mull on final project</p> <p>Develop an idea for a final project and post in tracker</p>

course	date	wk no.	session links	learning objectives	out-of-class work
DAT-102	Sat 24-APR'21	11	<p><i>Recording 1: Will Walker</i></p>  <p><i>Recording 2: Group ex review</i></p>  <p>passcode: %t3chnology%</p> <p>US Cesus and ACS</p> <p>Guest Analyst: William Walker</p> <p>The longest-running and most comprehensive sample-based data set is the US Census American Community Survey (ACS), the data from which is publicly accessible and incredibly rich.</p> <ul style="list-style-type: none">  American factfinder home  Site analysis GSheet  American Community Survey Error Rates Explained <p style="text-align: right;"> archive</p>	<p>TR.102.DS.7.A - Experiments - Designing:</p> <p>TR.102.DS.7.B - Experiments - Treatment assignment & Implementing:</p> <p>TR.102.DS.7.C - Experiments - Analyzing:</p> <p>TR.102.Q.10 - Standard errors</p> <p>TR.102.Q.11 - Student's T-tests - Setup</p> <p>TR.102.Q.12 - Student's T-tests - Interpretation</p>	<p>Dig into the Opp Atlas</p> <p>Please complete the exercises 0 and 1 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.</p> <p><i>Est. Time: 3-ish hours</i></p> <p>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.</p> <p><i>Start thinking about your final project</i></p> <p> DAT-102 Final project specs</p>
DAT-102	Sat 1-MAY'21	12	<p> Screen cast of class session CI review; opp atals overview</p> <p>Opportunity Atlas mini-project: multi-type data policy inquiry</p>	1	<p>Begin final project</p> <ul style="list-style-type: none">  DAT-102 Final project Board layout  DAT-102 Final project specs  DAT-102 Final project Group Tracker

course	date	wk no.	 Exploring the Opportunity Atlas <small>session links</small>	learning objectives	out-of-class work
			<p>Opp Atlas 2</p> <p> Data-based program evaluation</p> <p>Final project practice and design</p> <p> DAT-102 Final project specs</p>	<div style="border: 1px solid black; padding: 5px; display: inline-block; margin: 10px auto;">  archive </div>	<p>OPTIONAL Out of class: <i>Digest PGH Inequality report</i></p> <p>Due to COVID-19 reorganization, 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:</p> <ol style="list-style-type: none"> 1. Review the study's aggregation of 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 categories? 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 paragraph which starts: "When outcomes, like grade retention 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?

course	date	wk no.	session links	learning objectives	out-of-class work
DAT-102	Sat 8-MAY'21	13	<p><i>FA20: Session Recording</i></p>  <p> Experimental design</p> <p>Final project concept development</p> <p> archive</p>	1	1
DAT-102	Sat 15-MAY'21	14	<p>FINAL EXAM PERIOD from 10:00am- 12:00noon</p> <ul style="list-style-type: none">  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 <p> archive</p>		

Page created in 2021 by Eric Darsow and all non-linked content can be freely reproduced without any permission or attribution according to the [site's content use agreement](#). Any links to other content is governed by each page's respective usage rights context.