FALL 2022 // DATA 151
Course Calendar

Subject to Change

Course: Willamette University - DATA 151 - Introduction to Data Science		
Term: FALL 2022		**Assignments**
1A: Aug 30 Topics: ● Syllabus, Introductions, and Community Building ● Service available in accessible support *Motivating Questions: What is data science? What does data science mean to you?* **Related Reading:** ● *United Nations: Big Data for Sustainable Development* (link)	**1B: Sept 1** Topics: ● The Power of Data ● Types of variables *Motivating Questions: What is data? Where does data come from? How is data used for good?* **Related Reading:** ● iMStat - Ch 1 ○ 1: Hello Data	**Tasks: (Due 9/8)** ● Sign up to DataCamp ● Student Survey **HW #1: (Due 9/8)** ● Dear Data **Project: (Due 9/8)** ● Complete project partner survey
2A: Sept 6 Topics: ● Sampling Principles and Strategies **Related Reading:** ● iMStat - Ch 2: Study Design Sections: ○ 2.1: Sampling Principles and Strategies	**2B: Sept 8** Topics: ● Experiments and Principles of Experimental Design **Related Reading:** ● iMStat - Ch 2: Study Design Sections: ○ 2.2: Experiments ○ 2.3 Observational Studies	**HW #2: (Due 9/15)** ● Practice problems **Project Milestone #0: (Due 9/15)** ● Partner communication plan
3A: Sept 13 Topics: ● Introduction to R and *How to use the R Studio Cloud* ● *Introduction to R Markdown* ● Basics of R (Part I)	**3B: Sept 15** Topics: ● Basics of R (Part II) *Conditionals (if statements), types of loops, using simple functions, writing functions*	**HW#3: (Due 9/22)** ● DC: Introduction to R **Project Milestone #1: (Due 9/22)** ● Project proposal - include three

(FALL22) DATA 151: Course Calendar

● iDatSci : Ch 1 and 2 ○ Ch 1: Getting started with R and RStudio ○ Ch 2: R basics	basics	
4A: Sept 20 Topics: ● Data Wrangling ● Tidyverse (Part 1) - dplyr verbs **Related Reading:** ● iDatSci - Ch 4 ○ Ch 4: The tidyverse	**4B: Sept 22** Topics: ● Data Wrangling ● Tidyverse (Part 2) - tidyr **Related Reading:** ● iDatSci - Ch 4 ○ Ch 4: The tidyverse	**HW #4: (Due 9/29)** ● DC: tidyverse **Project Milestone #2:** ● Meet with Professor Smalley
5A: Sept 27 Topics: ● Importing Data ● *Review for Midterm* **Related Reading:** ● iDatSci - Ch 5 ○ Ch 5: Importing data	**5B: Sept 29** **MIDTERM #1** **R SKILLS CHECK**	**HW #5: (Due 10/6)** ● DC: Importing Data in R **Project Milestone #2:** *(if didn't meet week 4)* ● Meet with Professor Smalley
6A: Oct 4 Topics: ● Introduction to data visualization ● ggplot2: The layered grammar of graphics **Related Reading**: ● iDatSci - Ch 7 and 8 ○ Ch 7: Introduction to data visualization ○ Ch 8: ggplot2	**6B: Oct 6** Topics: ● What is the exploratory data analysis (EDA) process? **Related Reading:** ● R4DS: ○ Ch 7: Exploratory Data Analysis	**HW #6: (Due 10/13)** ● DC: Introduction to Data Visualization **Project Milestone #3: (Due 10/13)** ● EDA Step 1: Ask questions and form hypotheses
7A: Oct 11 Topics: ● EDA for categorical data ○ Simple bar graphs ○ Pie charts **Related Reading**:	**7B: Oct 13** Topics: ● EDA for categorical data ○ Tables and types of distributions and	**HW #7: (Due 10/20)** ● DC: Exploratory Data Analysis with Categorical Data

(FALL22) DATA 151: Course Calendar

Updated automatically every 5 minutes

Topics: ● EDA for numeric data ○ Histograms ○ Density plots ● Describing numeric distributions ○ Mean ○ Variance / standard deviation **Related Reading**: ● iMStat: Ch 5 ○ Exploring numerical data	Topics: ● EDA for numeric data ○ Box Plots ○ Exploring subgroups for comparison ● Robust metrics ○ Median ○ IVR ● Defining Outliers **Related Reading**: ● iDatSci: Ch 12 ○ Robust Summaries	● DC: Exploratory Data Analysis with Numerical Data **Project Milestone #5: (Due 10/27)** ● EDA Step #3: Distributions, Summary statistics, and Comparing subgroups
9A: Oct 25 Topics: ● Fun bonus topic! ○ Exploring spatial data ○ Advanced graphics	**9B: Oct 27** Topics: ● Critical thinking about graphics ● Best practice **Related Reading:** ● iDatSci: Ch 11 ○ Data visualization principles	**HW #9: (Due 11/3)** ● DC: Exploratory Data Analysis with
10A: Nov 1 *Review for Midterm*	**10B: Nov 3** **MIDTERM #2** **DATA VIZ and DESCRIPTIVE STATISTICS**	
11A: Nov 8 Topics: ● Two Dimensional Relationships between variables **Related Reading**: ● iMStat: Ch 7 Section ○ 7.1 Fitting a line, residuals, and correlation	**11B: Nov 10** Topics: ● Introduction to simple linear regression **Related Reading**: ● iMStat: Ch 7 Sections ○ 7.2 Fitting a line, residuals, and correlation ○ 7.3 Outliers in linear regression	**HW #10: (Due 11/17)** ● DC: Correlation and Regression **Project Milestone #6: (Due 11/17)** ● Create a scatter plot and describe the relationship between two numeric variables
12A: Nov 15 Topics: ● Communicating Statistical Results ● Creating a narrative with data	**12B: Nov 17** Topics: ● Pulling it all together and preparing for inference	WORK ON PROJECT PRESENTATION

(FALL22) DATA 151: Course Calendar

Project Presentations Day 2	Project Presentations Day 3	

Final Write-up in lieu of final exam.
Due during the assigned final time by the registrar.