DATA SCIENCE

Naumaan Nayyar

LEARNING OBJECTIVES

‣ Data Visualization

‣ Using Git to create and share your work

‣ Performing exploratory analysis - full stack!

‣ Describe Probability vs Odds

PRE-WORK

PRE-WORK REVIEW

‣ Explain the difference between variance and bias

‣ Use descriptive stats to understand your data

DATA VISUALIZATION

LET'S DISCUSS THE CURRENT LESSON OBJECTIVES

‣ Visualizing data is very important

‣ Let's spend some time going through a Data Visualization lab!

INTRODUCTION

GIT

GIT

‣ Version control is necessary when working on complex projects.

‣ Git is a way of tracking changes we've made to our programs that allows us to go back in time to fix errors.

‣ Combined with Github, Git is a powerful tool for collaborating with colleagues. You can work on different aspects of projects simultaneously and merge the changes together seamlessly.

‣ There are many different ways to use these tools.

GIT

‣ Let's see an example of using Git and Github.

‣ There are three primary commands we'll use.

    ‣ `git add`

    ‣ `git commit`

    ‣ `git push`

‣ When we want to implement a colleague's change, we may use the command `git pull`.

GIT

ACTIVITY: GIT

EXERCISE

DIRECTIONS (20 minutes)

1. Let's set up student repositories and move project and student work to them.

DELIVERABLE

Questions

EXPLORATORY DATA ANALYSIS

ACTIVITY: GIT

EXERCISE

DIRECTIONS (45 minutes)

1. Let's work on a dataset! http://archive.ics.uci.edu/ml/datasets/Adult

DELIVERABLE

Questions

ODDS AND PROBABILITY

ACTIVITY: ODDS & PROBABILITY

EXERCISE

DIRECTIONS (20 minutes)

Some of you may already be familiar with odds and probability.

1. We will use the starter code in lesson-05 of the class repo to review the concepts of odds and probability.

DELIVERABLE

Answer the questions in the notebook

TOPIC REVIEW

REVIEW

▸ Why is data visualization important?

▸ Are you comfortable with Git?

▸ Are you comfortable with exploratory analysis now…?

Q & A

EXIT TICKET

DON'T FORGET TO FILL OUT YOUR EXIT TICKET