**Data Analytics Boot Camp**

**Notes**

**Cheet sheet**

* **Commonly used terminal commands**
* cd (changes directory)
* cd ~ (changes to home directory)
* cd .. (moves up one directory)
* ls (lists files in folder)
* pwd (shows current directory)
* mkdir <FOLDERNAME> (creates new directory)
* touch <FILENAME> (creates a file)
* open . (opens the current folder. MAC SPECIFIC) or explorer . (opens the specific file. WINDOWS SPECIFIC)

**Side Notes:**

1. Use the tab key to auto-complete that file or directory name.
2. Auto-completion only works when we have a unique file or directory name that matches.
3. When naming your directories and files, we generally want to avoid having spaces in the name. If you wish to have a descriptive name, it's recommended you use underscores. So, instead of naming a directory terminal prework, name it terminal\_prework

**Analytics Paradigm - Steps on how to identify data “data problem solving”**

Step 1: Decompose the ask (i.e. what is an American? How to categorize food?)

Step 2: Identify Data Sources (i.e. don’t poll if there are platforms with this info)

Step 3: Define Strategy and Metrics (i.e. Americas will be based on cities with highest population)

Step 4: Build Data Retrieval Plan (i.e. use API’s, then build python script to randomly select 700 zip codes from US Census)

Step 5: Retrieve Data

Step 6: Assemble and clean the data

Step 7: Analyze for Trends (i.e. table, ratings by charts, statistical analysis such as “t-test”

* Limitations of Analysis

Step 8: Make the Call or tell the story (conclusion paragraph that may include limiation)

**EXCEL (Class #2)**

Example: SUM ( 1,2,3 )

-SUM = Functions

-123=Arguments

-Parameters = )

-Variable = what is the cell or what you renamed the cell

-Value = is the information in the cell

-Condition = Directions

Excel Tips:

Title cell or multiple cells by renaming it



When summing or formatting multiple sheets from an excel book use “!”

Vlookup is putting two tables together to have it meet in one column

False = exact match

True = estimated match