***Udacity Data Analyst Track***

**I. Into to Descriptive Stats**

1. Into to Research Methods

* **Constructs 🡪** anything that is difficult to measure, as it/they can be defined and measured in many different ways
* memory, humor
* intelligence 🡪 IQ test, grades, SAT’s
* effort 🡪 minutes doing homework, grades, GPA
* hunger 🡪 how often tummy grumbles, how nutrient-deficient one is
* itchiness 🡪 how many itchy areas, how often you’re itchy
* age
* volume **🡪** know it’s space something takes up but we have not yet defined how we are measuring that space. (i.e. liters, gallons, etc.)
* **Operational Definition (of a construct) 🡪** the unit of measurement used for the construct.
* *Once we operationally define something it is no longer a construct.*
* Had we said volume *in liters*, this would not be a construct because now it is operationally defined
* Minutes is already operationally defined; there is no ambiguity in what we are measuring
* **Population** = all the individuals in a group.
* **Sample** = *some* of the individuals in a group.
* **Parameter =** defines a characteristic of the population
* mean of a population is defined with the symbol **µ**
* **Statistic =** a characteristic of the sample.
* mean of a sample is defined as **¯x**
* **Sampling Error 🡪** difference between a parameter and a statistic
* **Treatment 🡪** the manner in an experiment in which researchers handle subjects
* specifically interested in how different treatments might yield differing results.
* **Observational Study 🡪** experimenter watches a group of subjects and does not introduce a treatment (i.e. survey)
* Only for correlation/RELATIONSHIPS
* Benefits of surveys:
* Easy way to get info on a pop (some can still be a bit difficult), relatively inexpensive, can be conducted remotely (web and mail), anyone can access and analyze results
* Often used to analyze constructs, since there isn’t just one definition
* Important to be as objective as possible and carefully though-out and worded
* Downsides of surveys:
* Untruthful and/or biased responses, not understanding the question (**response bias)**, refusing to answer **(non-response bias)**
* **Controlled Experiment 🡪** experimenter introduces a treatment/intervention
* To show **CAUSATION**

**Independent Variable 🡪** variable experimenters choose to manipulate (x-axis of a graph)

* **Dependent Variable 🡪** variable experimenters choose to measure during an experiment (y-axis)
* **Extraneous/Lurking Factors Variables** 🡪 variables not accounted for that affect the relationship between the IV(‘s) and the DV
* **Treatment Group** 🡪 receives varying levels of the independent variable
* used to measure the *effect* of a treatment.
* **Control Group 🡪** receives no treatment and is used as a baseline when comparing treatment groups
* **Placebo** 🡪 Something given to subjects in the control group so they think they are getting the treatment, when, in reality, they are getting something that causes no effect to them. (e.g. sugar pill)
* **Blinding 🡪** technique used to reduce bias.
* **Double blinding** ensures both those administering treatments and those receiving treatments do not know who is receiving which treatment.