**March 20, 2018**

**What is Data Science?**

**March 22, 2018**

Ask questions with “SMART goals framework”

S – Specific

M – Measureable

A – Attainable

R – Reproducible/Transparent

T – Time-bound

Data types matter:

1. Cross Sectional vs longitudinal

Cross Sectional – no distinction between exposure and outcome; can’t talk about causation; hard to separate cause and effect; variables with long duration are over represented

Strengths cross sectional – often population based, generalizabilitiy, and reduced cost

Longitudinal Strengths – unambiguous temporal sequence, multiple outcomes can be measured

Weaknesses – expense, takes a long time to collect data, vulnerable to missing data