# Timeline:

* Oct 11, 2019:
  + Individual assignments:
    - Uncleaned data OR
    - Data source and recommended methods
* Oct 09, 2019+ - Misc
  + Jai / Tommy: Discuss DB and rankings
  + Nich: Recommend tables

# Initial thoughts on Source:

## Linkedin:

* Recently loosened scraping policies
* Has proven / peer verified Job titles
  + Can we grade by Senior/Jr/Company
  + What is the difference between endorsements and skills

## Job postings:

* Current/modern/timely job requirements
* Has proven / company verified Job titles
  + Can we grade by Senior/Jr/Company
* Can sometimes be too general / not accurate / not technical

## Articles:

* Top ten articles / interviews with actual data scientists / other ideas?
* Article on indeed

# Methods:

* How to identify a skill?
  + Negative / stopword list?
  + Frequency?

# Store Data:

* What questions should we ask?
* How does this influence our DB Schema?
* What queries / what kind of queries will we run?
  + Where are at primarily jobs located - highest by count, salary, various variant questions
  + How long are posting online for jobs before they are removed (hiredat primarily)