Notes on presentation prep:

Phil to present slides 1-9, timeframe 5-6 minutes (shoot for 5)

Emily to present on the tableau dashboard

1: Economic conditions and movie preference

2: Do economic conditions influence what type of movies we watch? We hypothesize that it does

3: Question going in:

* how to define economic environment?
* What types of media are we looking at? Data sources? Computational tools?
* RELEVANCE: who doesn’t c\care about the economy? Who doesn’t like movies? Let’s tie them together!

4: cycle back to raw data as necessary(e.g. smoothing)

5:

* inflation, market returns, unemployment, home buyer index, consumer sentiment.
* Adventure, comedy, action, drame, horror.
* Data: get into same time periods, e.g. repeat quarterly for 3 months. Excel to help shape data into charts, smooth revenue over a one year period.
* Normalized (mean=0, variance =1) before data learning.

6: time period: 2006-2020: saw some economic volatility in that time, so if there is a relationship, it should come through

7: some correlation between economic conditions, e.g. UM(university of Michigan) sentiment and homeindex are tightly related, may be overkill to include both

8: Machine learning model.

* Tried linear regression – nice when it works; ez to explain.
* Supervised Random Forest model worked better. Split data into training and test sets, graphed predicted vs. actual y\_test values.
* Random forest : less prone to overfitting, better at predicting outside a range, hard to explain.
* Shortcut: treated each genre as independent, even though all need to add up to 100%. **Y values are market share of each genre: e.g. what percentage of total movie receipts for that month belong to the genre.**

9: Liinear regression outputs – mostly not much better than drawing the averages

10: decision tree outputs: nice graphs, economic conditions explain about half the variance for action, less for others.

* Adventure – a bit of a mess – maybe try a different smoothing level?
* Handoff: Decision Tree outputs show some relationship, but it does not offer a clear explanation of the drivers.
* We used a graphical tool to see what warrants further investigation
* Emily!