

## Team 8: Project Pitch

**Project Pitch:** The overall goal of the project is to take the different data analysis and knowledge extraction tools that we cover in class and apply them to a real world data set to demonstrate application of the aforementioned tools. Specifically, our goal is to apply the tools to collect data about Agatha Christie's novels and predict the order of the murders in her novel "And Then There Were None"

### Specific Aims:

- Find a dataset
- Methods:
  - Neural Net?
  - Decision Trees?
  - Unsupervised clustering?
  - NLP?
- Final outcome: Develop some predictive models, learn some conclusions about the data and predict the order

### Timeline:

- Week 3: Hammer out project proposal
  - Brainstorm and select a dataset
  - Start collecting data
- Week 4:
  - Brainstorm and select 3 methods of knowledge extraction from the data
  - Continue finding data
- Week 5:
  - Pre-processing the data
  - Continue finding data
- Week 6:
  - Applying 1 method to the data
- Week 7:
  - Applying another method to the data
- Week 8:
  - Applying another method to the data
- Week 9:
  - Building the predictive model (or models) and finalising the outcome of the project

### Team Outline:

- Work together on choosing dataset
- Work together choosing methods and final outcome
- Each person chooses a model/method to apply to the data, tunes and tests it
- Each person develops a part of the final presentation regarding their model
- **Program Manager Assignments**
  - Yogi -Week 3, 4, 9 &10
  - Ash Parasa - Weeks 5 & 6
  - William Huang - Week 7 & 8

**Final Deliverable:**

- Jupyter Notebook with all methods and models