

## First Draft of Group Project Pitch/Summary

**Group Members:** Austin Bryant, Andrew Nguyen, Brandon Davis

**Summary:** For our project, we found a Kaggle dataset that contained COVID-19 data along with party affiliation, racial demographics and other data at the county level. In the NYT COVID-19 dataset, we found mask use/preference also broken down by county. Our goal with these datasets is to determine if we can accurately predict mask use by county given some combination of the features available.

### Datasets:

<https://github.com/nytimes/covid-19-data>

<https://www.kaggle.com/etsc9287/2020-general-election-polls>

### Objectives:

- Pitch - what is your big picture (with a clear goal KDD related goal)
  - Analyze publically available data on a county level to predict the probability of mask wearing
- Specific Aims - 3-5 specific questions and potential results
  - Does a county's political leaning accurately predict mask use, and to what degree?
  - Are the racial demographics of a county a good predictor for mask use?
  - Are the economic demographics of a county a good predictor for mask use?
- Timeline
  - Week 4: Data cleaning, joining, train/test split
  - Week 5: NN research, decide on model aspects
    - Deliverable: Final formatted dataset
  - Week 6: Construction of model
  - Week 7: Training model, tuning
    - Deliverable: Finished model
  - Week 8: Gather metrics (accuracy, importance, etc.)
  - Week 9: Presentation work, visualizations
    - Deliverable: Altair visualizations
  - Week 10: Presentation work
    - Deliverable: Powerpoint presentation of findings
- Team Outline (who does what)
  - Network building/training will be split among the members
  - Austin Bryant: Dataset formatting and model inputs
  - Brandon Davis: Data visualization and exploratory analysis
  - Andrew Nguyen: Model selection(s) and figuring out possible libraries to use (scikit-learn, Tensorflow, ...)
- Final Deliverable (website/presentation/report/etc).
  - .PPT presentation with all findings and metrics