Understanding *COVID* Transmissions

through people, policies and mobility.

Harris County (Houston, TX)

DATA 512 A6

By

Venkata Sai Muktevi

1. Survey Data qualitative data

- At the State Level (Texas)
- 1,483 data points (people) for Texas from April 2020.
- Survey Consists of:
  - **Employment** status, **Education** levels, Demographics
  - **Measure of intent to follow different policies** on scale of 0-100.
  - Coronavirus Concern – measured on a real-number scale of 0-10.

Research Question

Aim: **To understand what influences the levels of concern for covid among people.**

Is the Coronavirus Concern of a person influenced by their…

1. Education levels,
2. Employment status,
3. Measure of intent to follow basic covid policies.

Will focus on this.

Intent to Follow Policies

- Policies under consideration:
  - Masking, Stay at Home, Six Feet Distance, Washing Hands
- 0 not likely to follow policy, 100 most likely to follow policy.
- Performed a Multiple Linear Regression with coronavirus concern as the response and intent to follow each of these policies as predictors.

Intent to Follow Policies

- Shows significant influence on covid concern from only intent to follow the Masking, Six Feet Distance and Stay at Home policies.

	coef	std err	t	P>|t|	[0.025	0.975]
const	-4.337e-18	0.025	-1.76e-16	1.000	-0.048	0.048
coronavirusIntent_Mask	0.0567	0.027	2.085	0.037	0.003	0.110
coronavirusIntent_SixFeet	0.0990	0.029	3.452	0.001	0.043	0.155
coronavirusIntent_StayHome	0.2219	0.028	7.904	0.000	0.167	0.277
coronavirusIntent_WashHands	0.0302	0.027	1.113	0.266	-0.023	0.083

- Initial hypothesis about people with more intent to follow policies are more likely to be “covid concerned” holds true.

Findings

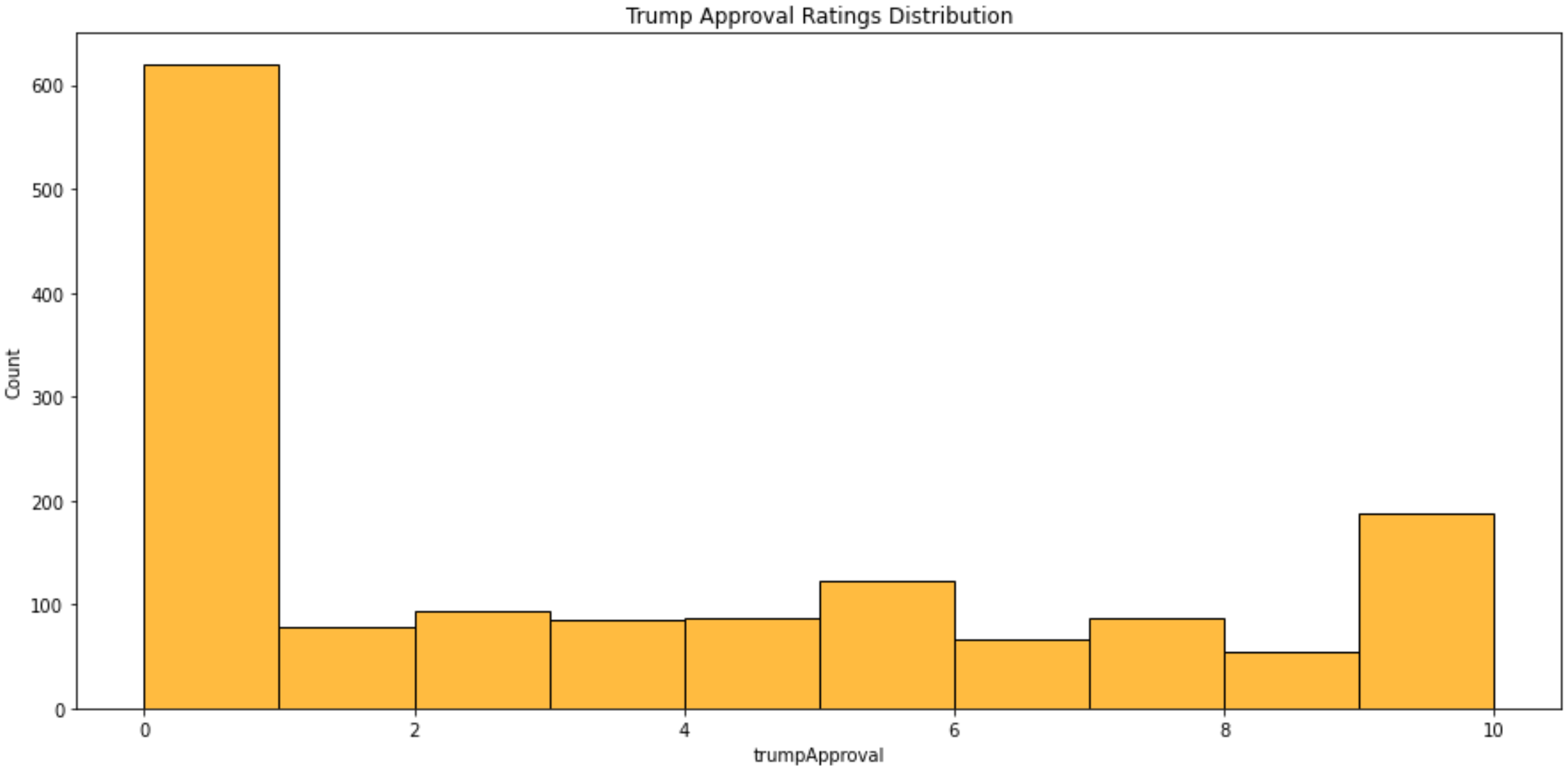
Gives us insight into what relates to people's covid concerns.

- Turns out that people with more concern are those that follow more extreme policies with more intent i.e, will Stay at Home rather than just only masking.

Also found that…

- Not enough evidence of influence on coronavirus concern based on employment status or education levels. (used ANOVA)

Intent to Follow Policies



2. Mobility Data quantitative data

- At the County Level (Harris County - Houston)
- 339 data points (days)
- Consists of:
  - Mobility data - A normalized version of the Number of times a type of route was requested on apple maps per day.
  - **Travel by Walking, Driving or Transit.**

Research Question

Aim: To understand if certain kinds of mobility correlate with covid transmissions indicated by the number of cases per day.

Do Covid cases per day show any meaningful relation with the various kinds of mobility data?

Analysis

- Performed a multiple linear regression with covid cases per day as the response and the various types of mobility as predictors.
- Showed some significant results in explaining direct or inverse relations with the target variable for Apple Mobility Data.

	coef	std err	t	P>|t|	[0.025	0.975]
const	1.583e-16	0.046	3.45e-15	1.000	-0.090	0.090
Apple_WalkingMobility	0.5338	0.136	3.923	0.000	0.266	0.801
Apple_TransitMobility	-0.4413	0.060	-7.348	0.000	-0.559	-0.323
Apple_DrivingMobility	-0.0081	0.150	-0.054	0.957	-0.304	0.287

Findings

Apple mobility data shows that **Covid cases per day** has:

- a significant *direct* relation with Walking route requests
- an *inverse* relation with Transit route requests.

Further research into hypotheses based on these relations can give us more information on transmission rates.

Issues, Pitfalls, Comments

- Size of data too small.
- Survey data is Ordinal.
- Survey data at the State level. No granularity.
- Mobility data biased to Apple users.
- Need further research and analysis with more robust testing.
- Some assumptions may be invalid due to lack of more information on data.
- Survey data can be considered ordinal.
- Survey data potential bias, either political or regional.
- Look into other tests or non-parametric tests to overcome issues in some analyses.
- Can’t make any causal conclusions only looking at meaningful relations.