



fig 1 "Water World"

Sip, Sip...Pass!

Data Analysis by Tatiana Thomas & Kai Hall

Topic: Drinking Water Standard Violations vs. Median Household Income

Description: This project analyzes the correlation between median household income and reported drinking water safety violations in the US countries in 2018.

Hypothesis: We hypothesize that there is a negative correlation between median household income and drinking water safety violations.

Null Hypothesis: There is not a significant correlation between median household income and drinking water safety violations.

Research Questions

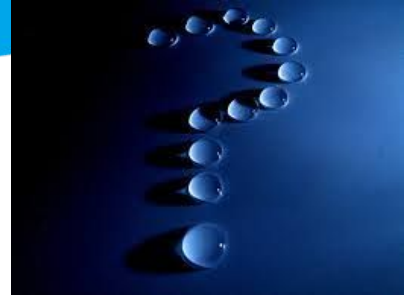


fig 2."Water Question"

1

Does income impact adherence to drinking water safety regulations?

2

What is the relationship between income and drinking water safety across United States countries?

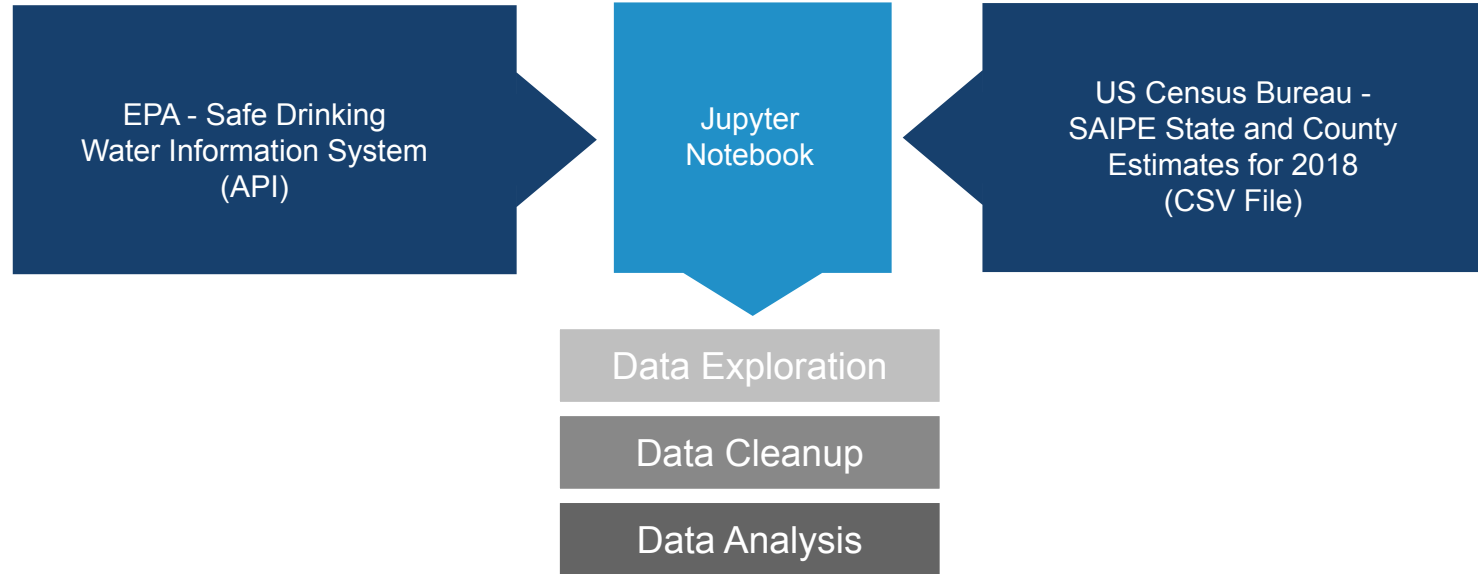
3

What are the top five chemicals referenced in the violations?

4

What are the major health effects of the top five chemicals?

Data Sets



Data Exploration and Cleanup

EPA Data (Tatiana) - Over 100,000 lines of data

- Exclude lines without counties
- By county
 - Total Number of Safety Violations
 - Chemical Name (SDW_VIOL_ENFORCEMENT.CNAME)
 - Violation Source (SDW_VIOL_ENFORCEMENT.SOURCES)
 - Definition (SDW_VIOL_ENFORCEMENT.DEFINITION)
 - Health Effects (SDW_VIOL_ENFORCEMENT.HEALTH_EFFECTS)

Census Data (Kai) - 3,914 lines of data

- Exclude unnecessary rows (Rows 1-3)
- State (Postal Code)
- County (Name)
 - Removal of rows with aggregated State data
- Median Household Income (Column W)

Data Cleanup

- Merging on County Data
- Data format was inconsistent and county names differed from our original assumptions

```
# Identification of naming conventions outside of "County"
```

```
bad_census = new_census[new_census['Name'].str.contains("County")==False]  
bad_census
```

```
# Extracting rows with county-level data
```

```
clean_census = new_census[new_census['Name'].str.contains("County|city|Borough|Census|Parish|District")]  
clean_census.head()
```

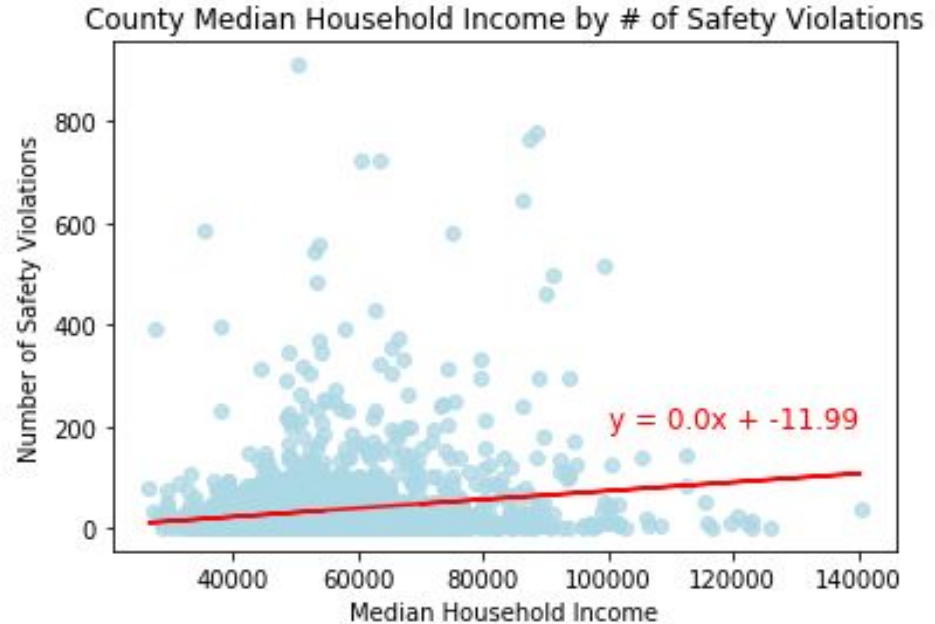
```
# Reformatting County data to match format of EPA data for merge
```

```
clean_census['Name2'] = clean_census['Name'].str.replace(' County', '')  
clean_census['NameUp'] = clean_census['Name2'].str.upper()  
  
clean_census.head()
```

RQ1: What, if any, impact does income have on the adherence to drinking water safety regulations?

The r-squared value is: 0.027

This indicates a weak or non-existent correlation between the median household income and number of drinking water safety violations.



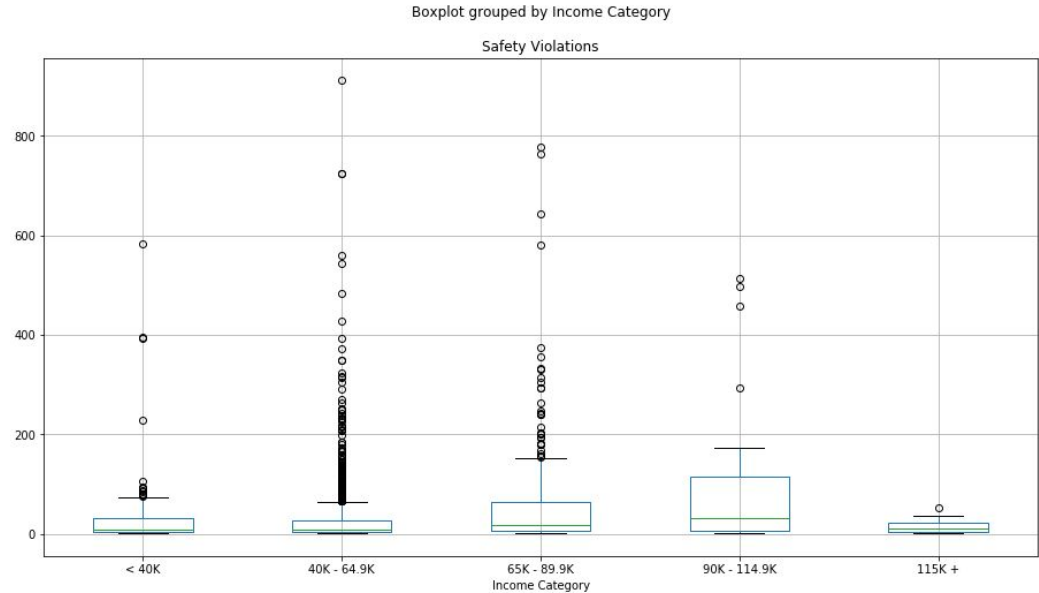
RQ1: What, if any, impact does income have on the adherence to drinking water safety regulations?

One-way ANOVA Test

p-value =

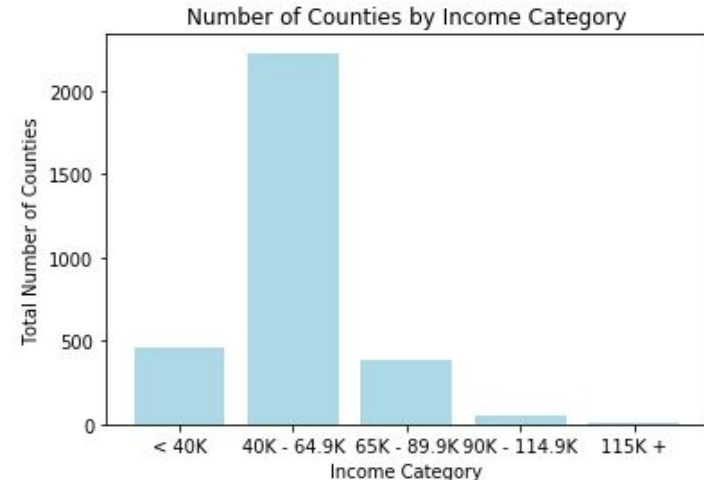
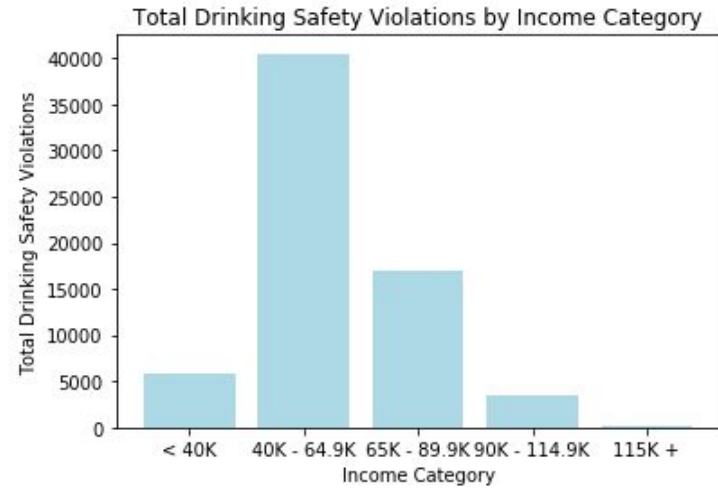
3.151998268007751e-12

This indicates that there is not a significant difference between the groups.



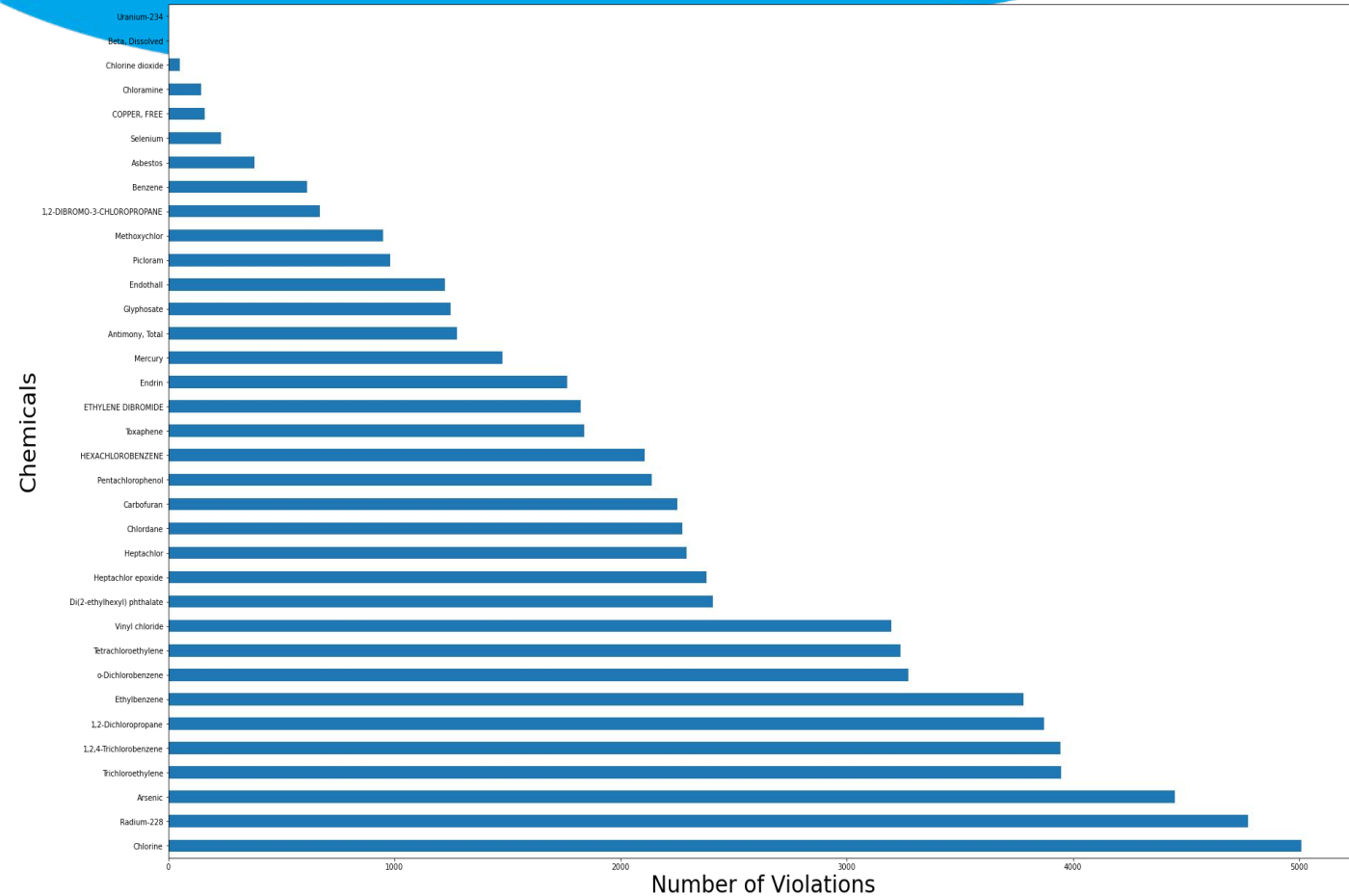
RQ2: What is the relationship between income and drinking water safety across United States counties?

The distribution of drinking water safety violations by income category is similar to the distribution of counties represented by income category.



Leading Chemicals of Violations

Number of Violations vs Chemicals



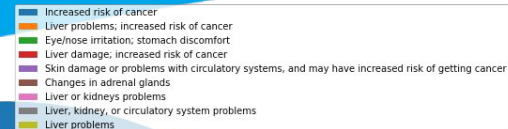
Of the 70,176 violations the top five chemicals involved are:

Chlorine
Radium-228
Arsenic
Trichloroethylene
1,2,4-Trichlorobenzene

35 Chemicals Total

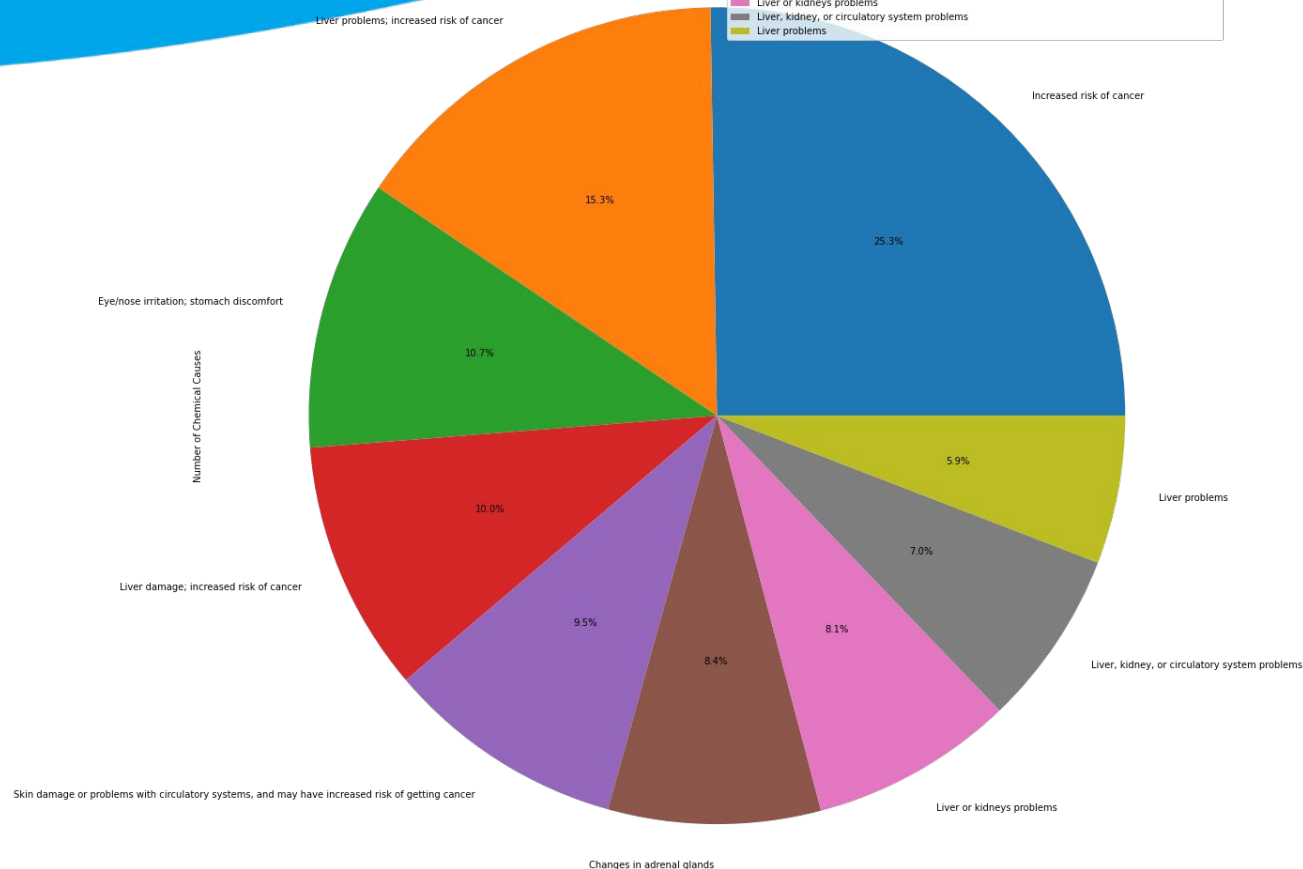
Major Health Effects

Major Health Effects

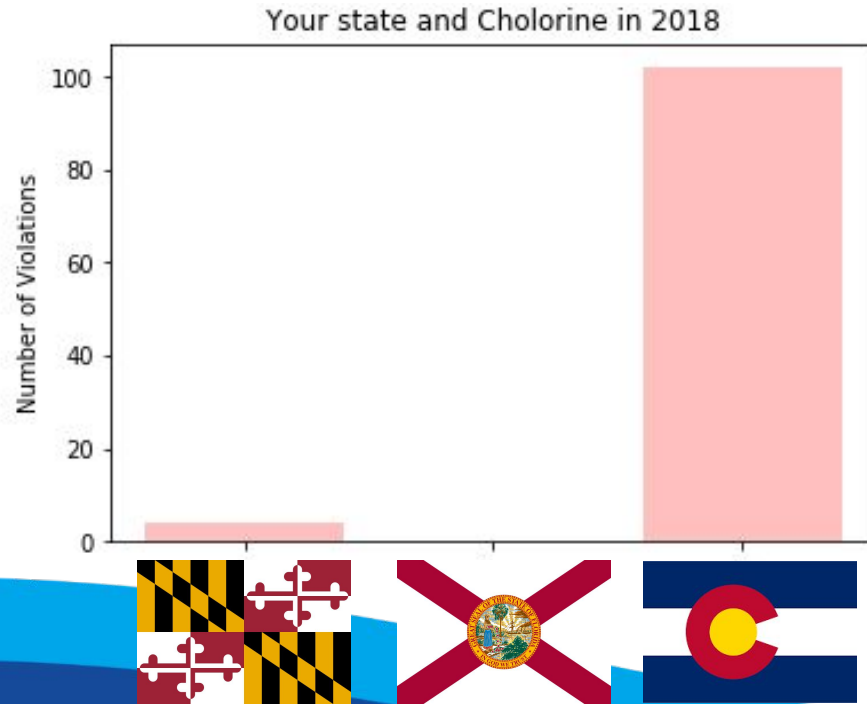


Leading Health Effects

- Increased risk of cancer
25.3%
- Kidney/Liver problems
15.3%
- Eye/nose
irritation/Stomach
discomfort 10.7%



Can you sip with a peace of mind?



[fig.3 "MD fla"](#)



[fig.4 "FL flag"](#)



[fig.5 "CO flag"](#)

Conclusions

Top Five Chemical Contaminants:

- Chlorine
- Radium-228
- Arsenic
- Trichloroethylene
- 1,2,4-Trichlorobenzene

Most Prevalent Potential Adverse Health Effects:

- Increased Risk of Cancer
- Kidney/Liver Problems
- Eye/Nose Irritation/Stomach Discomfort

Hypothesis:

We hypothesize that there is a negative correlation between median household income and drinking water safety violations.

Null Hypothesis:

There is not a significant correlation between median household income and drinking water safety violations.

We fail to reject the null hypothesis.

R-Squared:	0.027
P-Value:	3.151998268007751e-12

Implications and Opportunities

Implications

1. There is not a significant difference in the number of drinking water safety violations across income levels.

Opportunities

1. Impact analysis of most prevalent contaminants on the communities that are most hard-hit.
2. Analysis of the availability of water treatment and additional safety interventions in areas with the most water safety violations.
3. Global comparison of water safety violation measurement and income.



THANK YOU

Resources

Google Slides

- Select slide and edit your presentation
- You can download and share your presentation!

Powerpoint Templates

- Free Powerpoint Presentation Template (.pptx)
- You can upload Google Slides..

Free Powerpoint Templates:
Abstract Google Slides:
Animal and Wildlife Presentation:
Biology Powerpoint Templates:
Business & Finance Google Slides:
Chemistry Powerpoint Slides:
Christmas Google Slides:
Education Google Slides:
Flowers Powerpoint Templates:
Food and Drinks Slides:
Games Powerpoint Templates:
Medical Google Slides:
Nature Powerpoint Templates:
Religious Google Slides:
Sports Powerpoint Slides:
Technology Google Slides:
Transportation Powerpoint Templates:

<https://googleslides.org/free-templates>
<https://googleslides.org/free-templates/abstract-google-slides>
<https://googleslides.org/free-templates/animal-wildlife-presentation-template>
<https://googleslides.org/free-templates/biology-presentation-slides>
<https://googleslides.org/free-templates/business-finance-templates>
<https://googleslides.org/free-templates/chemistry-google-slides-template>
<https://googleslides.org/free-templates/christmas-new-year-templates>
<https://googleslides.org/free-templates/education-google-slides>
<https://googleslides.org/free-templates/flowers-powerpoint-templates>
<https://googleslides.org/free-templates/food-drinks-powerpoint-templates>
<https://googleslides.org/free-templates/games-powerpoint-templates>
<https://googleslides.org/free-templates/medical-powerpoint-templates>
<https://googleslides.org/free-templates/nature-google-slide-templates>
<https://googleslides.org/free-templates/religion-powerpoint-templates>
<https://googleslides.org/free-templates/sports-presentation-slides>
<https://googleslides.org/free-templates/technology-powerpoint-templates>
<https://googleslides.org/free-templates/transportation-google-slides>