

The background of the slide is a stylized American flag. The top-left corner features a blue field with white stars, while the rest of the slide is filled with horizontal red and white stripes. A white silhouette of the United States map is centered on the slide, serving as a backdrop for the text.

# **USA Charity Data Exploration**

*Jack Vessa - Sep 2019*

# ***USA Charity Data Exploration***

## **Project Focus**

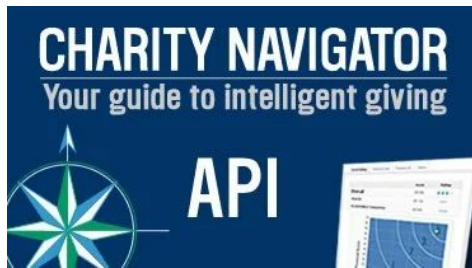
- Gather Information about USA Charities

## **Motivation**

- Interested in career making positive social impact
- Partnering with a charity donation organization

# ***Data Pipeline***

 **IRS**  
***Population***



***Sample (0.65%)***

python  
+  
**Pandas**



**matplotlib**



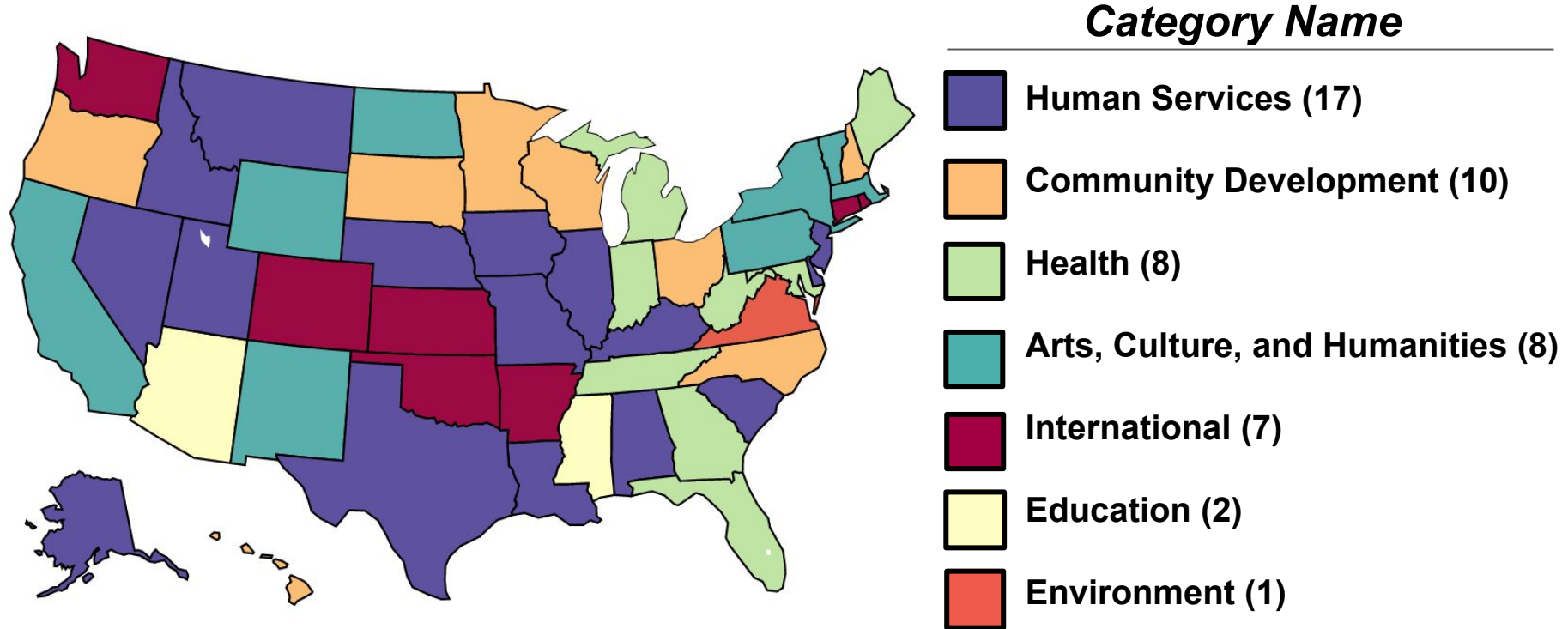
**SciPy**

# What is the *Mission* of USA Charities?



# What Does Each *State* Care Most About?

Investigating Yearly Donation Totals By Category



*\* includes DC, Puerto Rico & Virgin Islands*

# Let's Take a Look at Charity Region!

*Does Region affect  
AVG Charity Income?*



# ***United States Charitable Organizations by **Region** (8,642 Total)***

---

## **West Region (1,986)**



## **Midwest Region (1,615)**



## **Northeast Region (2,081)**



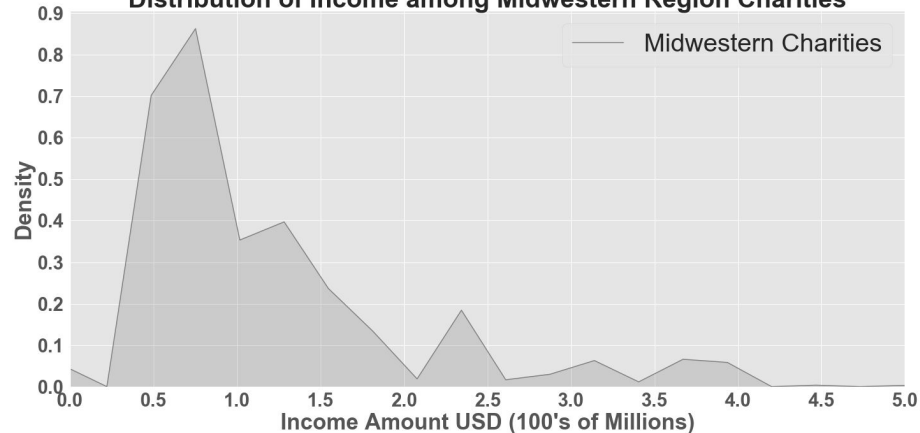
## **South Region (2,960)**



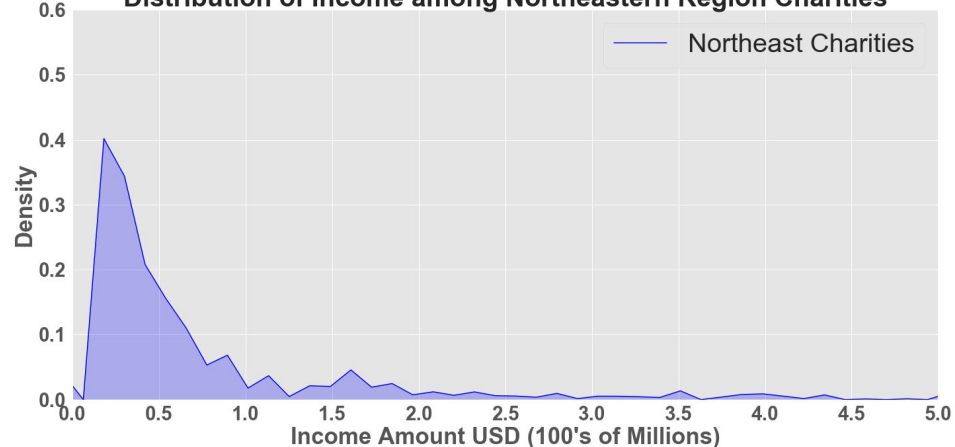
**Source:**  
[Business Insider](#)

# KDE of Average Charity Income by *Region*

## Distribution of Income among Midwestern Region Charities



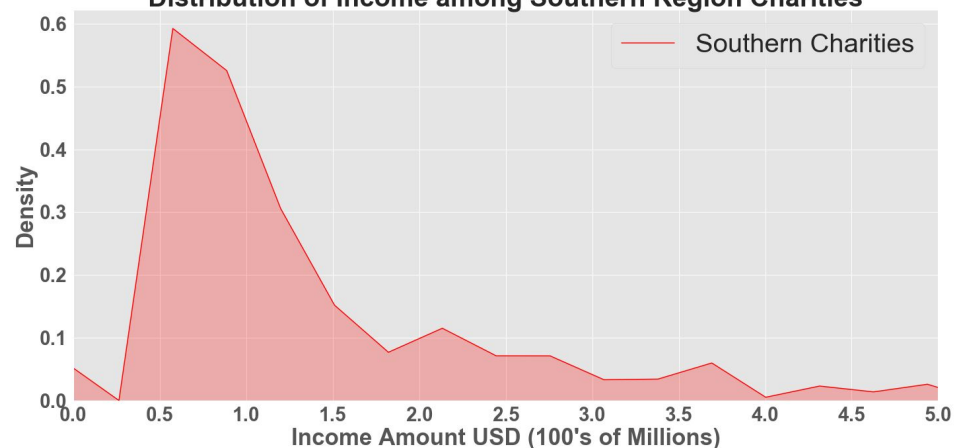
## Distribution of Income among Northeastern Region Charities



## Distribution of Income among Western Region Charities

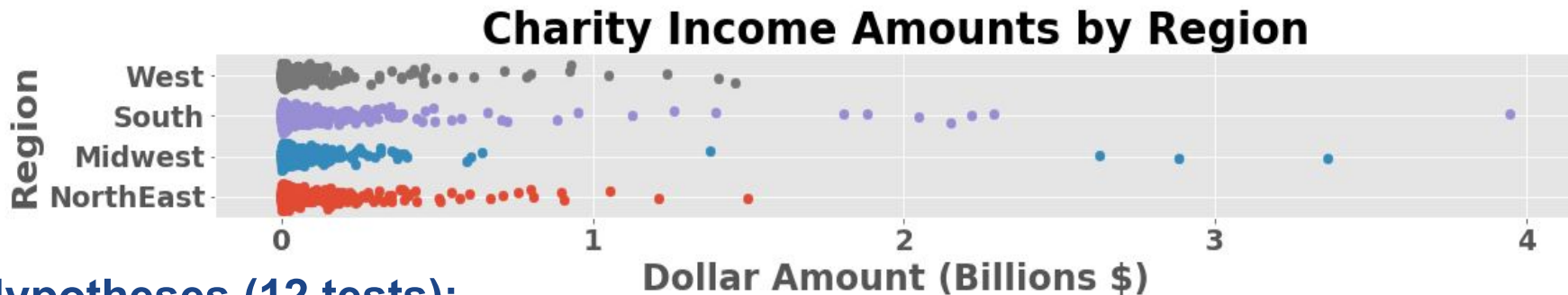


## Distribution of Income among Southern Region Charities





# Does *Region* Affect Average Charity Income?



## Hypotheses (12 tests):

$H_0 \rightarrow$  The average charity income in region 1 is less than/equal region 2

$\mu_1 \leq \mu_2$  (Compared across all regional pairs)

$H_A \rightarrow$  The average charity income in region 1 is greater than region 2

$\mu_1 > \mu_2$  (Compared across all regional pairs)

Significance Level:  $\alpha = 0.05$   
(Bonferroni Correction):  $\alpha = 0.05/12 = 0.004$

## Mann-Whitney Signed Rank Test

p-value for NorthEast  $\leq$  Midwest: 0.00012 \*

p-value for NorthEast  $\leq$  South : 0.00110 \*

p-value for NorthEast  $\leq$  West : 0.00583 \*

**Conclusion:** There is significant evidence to REJECT Null Hypothesis (NorthEast AVG charity income is higher than Midwest and South)

## Other Notable Results

p-value for West  $\leq$  Midwest : 0.10458 \*

p-value for South  $\leq$  Midwest : 0.14954 \*

# ***Project Takeaways***

- **Charities** care about **Providing Community Programs**
- **Northeast Charities** receive **MORE** donation income, on average, than **Midwestern** and **Southern** charities

## ***Future Exploration***

- Connect USA Charity Data to **Company User Data**
- Build **Charity Recommender** (Capstone 3)

# ***Thank You!***



# ***Questions?***

# Works Cited

- **Charity Navigator API Data.** Retrieved from:  
<https://www.charitynavigator.org/index.cfm?bay=content.view&cpid=1397>
- **IRS Charitable Organization Data.** Retrieved from:  
<https://www.irs.gov/charities-non-profits/exempt-organizations-business-master-file-extract-eo-bmf>
- **Business Insider USA Regional Graphs.** Retrieved from:  
<https://www.businessinsider.com/regions-of-united-states-2018-5#the-us-census-bureau-divides-the-united-states-into-four-regions-theres-the-north--1>