FreshSight Queen's

Donor Survey Analysis Model

September 10th, 2024

FreshSight requires a sophisticated, multi-phased tool to derive insights from its projected survey data

What are we hoping to learn from this data?

1. “What factors are most important to the stimulation of a donor base?
2. “What do our top donors look like? What about our bottom-level donors?
3. “How do different communities stack up in terms of donation magnitude and frequency?
4. “What can Non-Profits learn/about from their donor base?
5. “How can Non-Profits tailor their strategies to their donors?

Projected Analysis Structure

FreshSight requires a sophisticated, multi-phased tool to derive insights from its projected survey data

What are we hoping to learn from this data?

1. "What factors are most important to the stimulation of a donor base?
2. "What do our top donors look like? What about our bottom-level donors?
3. "How do different communities stack up in terms of donation magnitude and frequency?
4. "What can Non-Profits learn/about from their donor base?
5. "How can Non-Profits tailor their strategies to their donors?

Projected 3-Phased Analysis Approach

Multiple Linear Regression / Correlation

Implemented using a Scikit Linear Regression Model

Demographic Characteristic Analysis

Implemented using Pandas DataFrame manipulation

Qualitative Response Analysis

Implemented by manual manipulation

FreshSight requires a sophisticated, multi-phased tool to derive insights from its projected survey data

What are we hoping to learn from this data?

1. What factors are most important to the stimulation of a donor base?
2. What do our top donors look like? What about our bottom-level donors
3. How do different communities stack up in terms of donation magnitude and frequency?
4. What can Non-Profits learn/about from their donor base?
5. How can Non-Profits tailor their strategies to their donors?

Projected 3-Phased Analysis Approach

Multiple Linear Regression / Correlation

Implemented using a Scikit Linear Regression Model

Demographic Characteristic Analysis

Implemented using Pandas DataFrame manipulation

Qualitative Response Analysis

Implemented by manual manipulation

Multiple Linear Regression (MLR) is a statistical technique to model the relationship between a target variable and multiple predictors

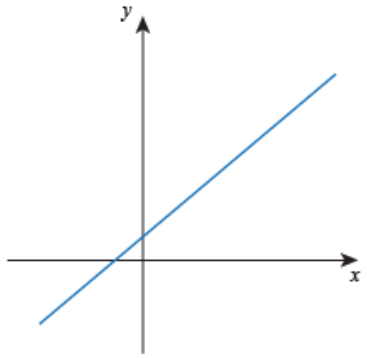
Brief Theory

The general form of an MLR model is as follows:

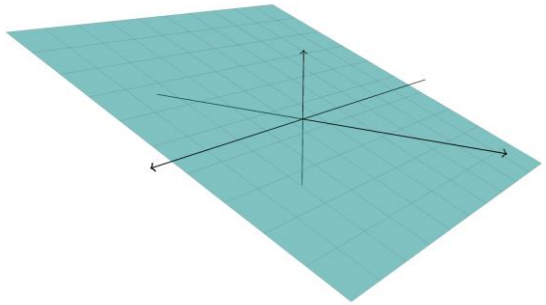
$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_n X_n + \epsilon$$

Where:

- **Y** is the dependent variable, referred to as the **target**,
- $\mathbf{X_i}$ are the independent variables, referred to as a **feature**,
- $\boldsymbol{\beta_0}$ is the intercept,
- $\boldsymbol{\beta_i}$ are the coefficients for each independent variable,
- $\boldsymbol{\epsilon}$ is the error residual.



For a single-featured function



For a double-featured function

What is Scikit-learn's MLR model?

1. Given a dataset consisting of target values in respect to multiple feature values, the LinearRegression() tool's goal is to **determine the coefficients and intercept of the MLR relationship** hosted within the dataset.

2. Based on these newly determine coefficients, the model can **predict the value of the target given new feature values.**

How will the model be used?

1. Here, our target variables will be average Donation Amount and Donation frequency. Our features will be some demographic data, and 1-10 scale answers regarding Non-Profit perception.

2. The model will be **trained on the dataset received from the survey**. From there, it can be loaded into a Python script to declare the coefficient values or predict new outputs.

FreshSight requires a sophisticated, multi-phased tool to derive insights from its projected survey data

What are we hoping to learn from this data?

1. What factors are most important to the stimulation of a donor base?
2. What do our top donors look like? What about our bottom-level donors
3. How do different communities stack up in terms of donation magnitude and frequency?
4. What can Non-Profits learn/about from their donor base?
5. How can Non-Profits tailor their strategies to their donors?

Projected 3-Phased Analysis Approach

Multiple Linear Regression / Correlation

Implemented using a Scikit Linear Regression Model

Demographic Characteristic Analysis

Implemented using Pandas DataFrame manipulation

Qualitative Response Analysis

Implemented by manual manipulation

Demographic Analyses will follow a Backwards-Forwards approach to provide a holistic view of community-based behaviour

The Backwards Approach

High

Low

Donation Amount	...	Char. 1	Char. 2	
XXXXX	...	Community A	Community B	Top 25%
XXXXX	...	Community B	Community C	
XXXX	...	Community A	Community C	
...	...	...	...	
XX	...	Community D	Community A	Bottom 25%
X	...	Community B	Community D	
X	...	Community C	Community E	

1. What are the communities with the highest representation in the Top/Bottom 25% of Donation Amount/Frequency?
2. What communities are underrepresented in these categories? Why?

The Forwards Approach

Donation Amount	...	Char. 1	Char. 2
XX	...	Community A	Community A
XXXXX	...	Community A	Community A
XXX	...	Community A	Community B
...	...	...	...
XXXX	...	Community D	Community E
XX	...	Community E	Community E
XXX	...	Community E	Community E

1. What is Community A's average Donation Amount/Frequency?
2. How does this stack up against Community B? Why is this the case?

FreshSight requires a sophisticated, multi-phased tool to derive insights from its projected survey data

What are we hoping to learn from this data?

1. "What factors are most important to the stimulation of a donor base?
2. "What do our top donors look like? What about our bottom-level donors
3. "How do different communities stack up in terms of donation magnitude and frequency?
4. "What can Non-Profits learn/about from their donor base?
5. "How can Non-Profits tailor their strategies to their donors?

Projected 3-Phased Analysis Approach

Multiple Linear Regression / Correlation

Implemented using a Scikit Linear Regression Model

Demographic Characteristic Analysis

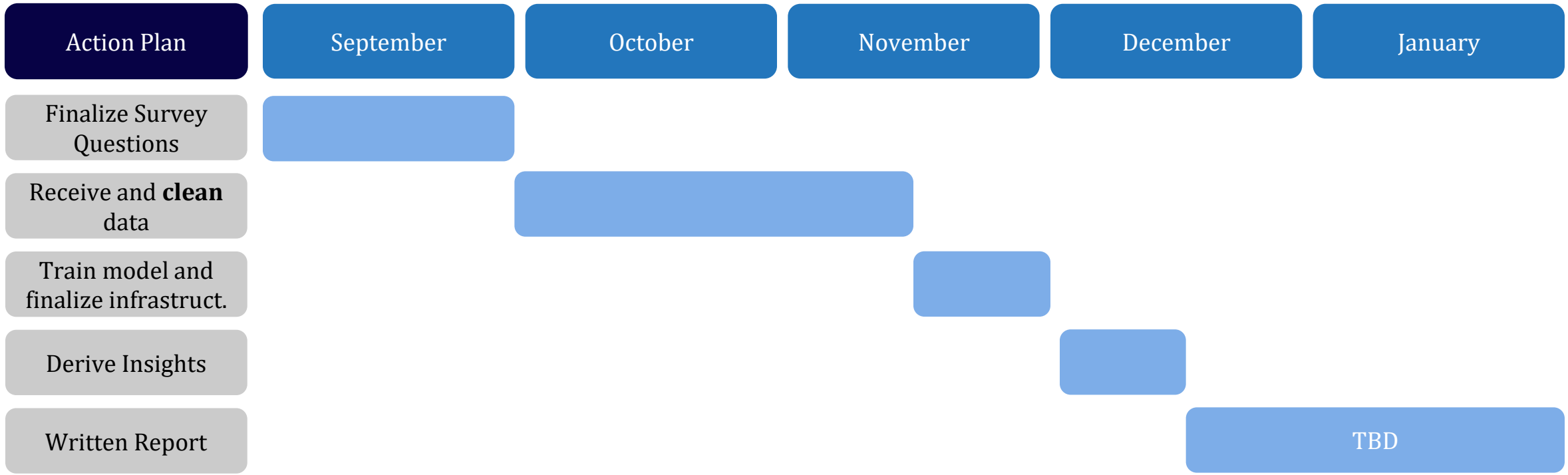
Implemented using Pandas DataFrame manipulation

Qualitative Response Analysis

Implemented by manual manipulation

Next Steps & Timeline

Survey questions and submission at forefront



The infrastructure is laid out for analysis. Our next steps involve working with the Research Impact Hub to set up the survey and using the model for a new publication!

Thank You

Questions