

US - Elections

Pramod  
Padmanabhan

Plan of the  
Talk

Introduction  
The Problem

Data and  
Variables

Visualizing the  
Dependencies

Inferences

Outlook

# Decoding Trump's Voters

Pramod Padmanabhan

23rd July, 2016

# Plan of the Talk

US - Elections

Pramod  
Padmanabhan

Plan of the  
Talk

Introduction  
The Problem

Data and  
Variables

Visualizing the  
Dependencies

Inferences

Outlook

## ① Introduction

# Plan of the Talk

US - Elections

Pramod  
Padmanabhan

Plan of the  
Talk

Introduction  
The Problem

Data and  
Variables

Visualizing the  
Dependencies

Inferences

Outlook

- ① Introduction
- ② The Predictors and the Predicted

# Plan of the Talk

US - Elections

Pramod  
Padmanabhan

Plan of the  
Talk

Introduction  
The Problem

Data and  
Variables

Visualizing the  
Dependencies

Inferences

Outlook

- ① Introduction
- ② The Predictors and the Predicted
- ③ Decision Trees

# Plan of the Talk

US - Elections

Pramod  
Padmanabhan

Plan of the  
Talk

Introduction  
The Problem

Data and  
Variables

Visualizing the  
Dependencies

Inferences

Outlook

- ① Introduction
- ② The Predictors and the Predicted
- ③ Decision Trees
- ④ Inferences

# Plan of the Talk

US - Elections

Pramod  
Padmanabhan

Plan of the  
Talk

Introduction  
The Problem

Data and  
Variables

Visualizing the  
Dependencies

Inferences

Outlook

- ① Introduction
- ② The Predictors and the Predicted
- ③ Decision Trees
- ④ Inferences
- ⑤ Outlook

# US Elections 2016

US - Elections

Pramod  
Padmanabhan

Plan of the  
Talk

**Introduction**  
The Problem

Data and  
Variables

Visualizing the  
Dependencies

Inferences

Outlook

- The Presidential elections in the United States is a quadrennial event that is keenly observed by the rest of the world.

# US Elections 2016

## US - Elections

Pramod  
Padmanabhan

Plan of the  
Talk

Introduction  
The Problem

Data and  
Variables

Visualizing the  
Dependencies

Inferences

Outlook

- The Presidential elections in the United States is a quadrennial event that is keenly observed by the rest of the world.
- The year 2016 gets marked by the nature of the candidates, Hillary Clinton, with the chance of becoming the first ever Woman President and the billionaire Donald Trump, for his extreme views.



# US Elections 2016

## US - Elections

Pramod  
Padmanabhan

Plan of the  
Talk

Introduction  
The Problem

Data and  
Variables

Visualizing the  
Dependencies

Inferences

Outlook

- The Presidential elections in the United States is a quadrennial event that is keenly observed by the rest of the world.
- The year 2016 gets marked by the nature of the candidates, Hillary Clinton, with the chance of becoming the first ever Woman President and the billionaire Donald Trump, for his extreme views.
- The Democrat, Bernie Sanders was a practical socialist who unfortunately did not get enough support needed for the nomination.

# The Problem - Trump's Voters

US - Elections

Pramod  
Padmanabhan

Plan of the  
Talk

Introduction  
**The Problem**

Data and  
Variables

Visualizing the  
Dependencies

Inferences

Outlook

- The world was taken by surprise when Donald Trump was nominated by the Republicans to contest in the general elections.

# The Problem - Trump's Voters

US - Elections

Pramod  
Padmanabhan

Plan of the  
Talk

Introduction  
**The Problem**

Data and  
Variables

Visualizing the  
Dependencies

Inferences

Outlook

- The world was taken by surprise when Donald Trump was nominated by the Republicans to contest in the general elections.
- We want to find out who are the people who voted for Trump giving him the ticket. Our goal is to describe these people.

# The Problem - Trump's Voters

US - Elections

Pramod  
Padmanabhan

Plan of the  
Talk

Introduction  
The Problem

Data and  
Variables

Visualizing the  
Dependencies

Inferences

Outlook

- The world was taken by surprise when Donald Trump was nominated by the Republicans to contest in the general elections.
- We want to find out who are the people who voted for Trump giving him the ticket. Our goal is to describe these people.
- We do this by looking at which racial group they belong to, their level of education, employment and poverty level, their profession, their income, if they belong to hate groups and if they come from urban or rural areas.

# The Predictors and the Predicted

US - Elections

Pramod  
Padmanabhan

Plan of the  
Talk

Introduction  
The Problem

Data and  
Variables

Visualizing the  
Dependencies

Inferences

Outlook

- The dependent variable is the number of votes gained by Trump at the Republican primaries held in the various states. We have the results from 46 such states, excluding Colorado, DC, Maine, Minnesota and North Dakota.

# The Predictors and the Predicted

US - Elections

Pramod  
Padmanabhan

Plan of the  
Talk

Introduction  
The Problem

Data and  
Variables

Visualizing the  
Dependencies

Inferences

Outlook

- The dependent variable is the number of votes gained by Trump at the Republican primaries held in the various states. We have the results from 46 such states, excluding Colorado, DC, Maine, Minnesota and North Dakota.
- The variables that the number of votes depend on, include the following :

# The Predictors and the Predicted

US - Elections

Pramod  
Padmanabhan

Plan of the  
Talk

Introduction  
The Problem

Data and  
Variables

Visualizing the  
Dependencies

Inferences

Outlook

- The dependent variable is the number of votes gained by Trump at the Republican primaries held in the various states. We have the results from 46 such states, excluding Colorado, DC, Maine, Minnesota and North Dakota.
- The variables that the number of votes depend on, include the following :  
Percentages of various racial groups in the state.

# The Predictors and the Predicted

US - Elections

Pramod  
Padmanabhan

Plan of the  
Talk

Introduction  
The Problem

Data and  
Variables

Visualizing the  
Dependencies

Inferences

Outlook

- The dependent variable is the number of votes gained by Trump at the Republican primaries held in the various states. We have the results from 46 such states, excluding Colorado, DC, Maine, Minnesota and North Dakota.

- The variables that the number of votes depend on, include the following :

Percentages of various racial groups in the state.

Economic indicators for the state described by the dominant revenue generator for the state.



# The Predictors and the Predicted

US - Elections

Pramod  
Padmanabhan

Plan of the  
Talk

Introduction  
The Problem

Data and  
Variables

Visualizing the  
Dependencies

Inferences

Outlook

- The dependent variable is the number of votes gained by Trump at the Republican primaries held in the various states. We have the results from 46 such states, excluding Colorado, DC, Maine, Minnesota and North Dakota.

- The variables that the number of votes depend on, include the following :

Percentages of various racial groups in the state.

Economic indicators for the state described by the dominant revenue generator for the state.

The amount of education, unemployment and poverty in a state.

# The Predictors contd ...

US - Elections

Pramod  
Padmanabhan

Plan of the  
Talk

Introduction  
The Problem

**Data and  
Variables**

Visualizing the  
Dependencies

Inferences

Outlook

The mean household income of each state.

# The Predictors contd ...

US - Elections

Pramod  
Padmanabhan

Plan of the  
Talk

Introduction  
The Problem

**Data and  
Variables**

Visualizing the  
Dependencies

Inferences

Outlook

The mean household income of each state.

Urbanization described by the number of metros and non-metro regions in each state.

# The Predictors contd ...

US - Elections

Pramod  
Padmanabhan

Plan of the  
Talk

Introduction  
The Problem

Data and  
Variables

Visualizing the  
Dependencies

Inferences

Outlook

The mean household income of each state.

Urbanization described by the number of metros and non-metro regions in each state.

Hate group indicator shown by the number of extremist organizations like the KKK in each state.

# The Predictors contd ...

US - Elections

Pramod  
Padmanabhan

Plan of the  
Talk

Introduction  
The Problem

Data and  
Variables

Visualizing the  
Dependencies

Inferences

Outlook

The mean household income of each state.

Urbanization described by the number of metros and non-metro regions in each state.

Hate group indicator shown by the number of extremist organizations like the KKK in each state.

- We use a total of 18 variables which are normalized to account for the different sizes of the states.

# The Predictors contd ...

US - Elections

Pramod  
Padmanabhan

Plan of the  
Talk

Introduction  
The Problem

Data and  
Variables

Visualizing the  
Dependencies

Inferences

Outlook

The mean household income of each state.

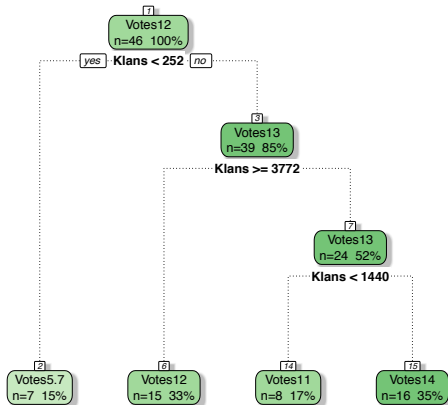
Urbanization described by the number of metros and non-metro regions in each state.

Hate group indicator shown by the number of extremist organizations like the KKK in each state.

- We use a total of 18 variables which are normalized to account for the different sizes of the states.
- Variables which we could have a significant effect but not used here include religious composition of each state, number of veterans in the state, immigrant make up of each state.

# Decision Trees

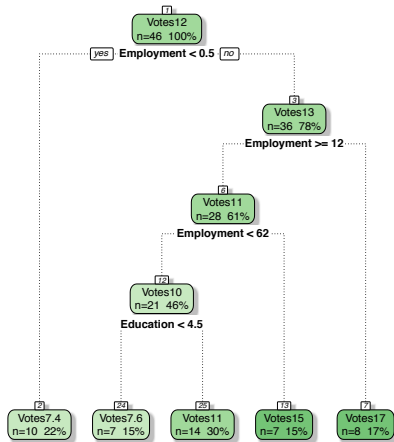
- Klan Dependence



Regression tree for Votes vs Klan Percentage, n denotes the number of states.

# Decision Trees contd...

- Education/Employment/Poverty

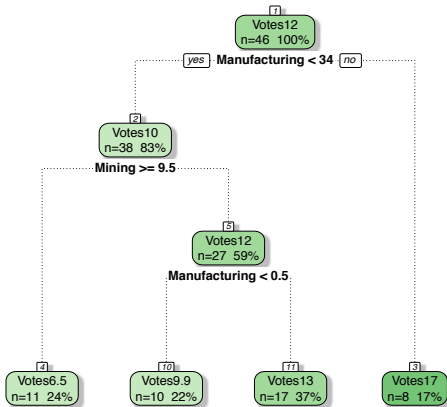


Regression tree for Votes vs Employment/Education/Poverty.



# Decision Trees contd...

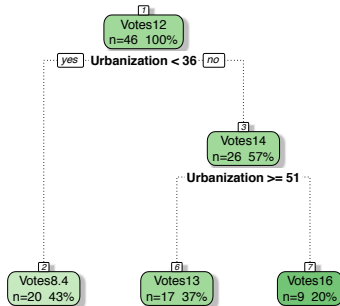
- Economic Dependence



Regression tree for Votes vs Type of Economy.

# Decision Trees contd....

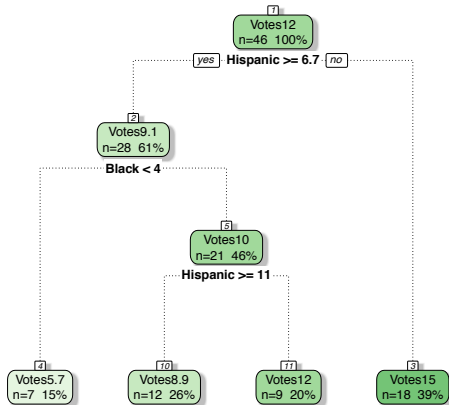
- Dependence of level of Urbanization of a state.



Regression tree for Votes vs Urbanization Percentage.

# Decision Trees contd....

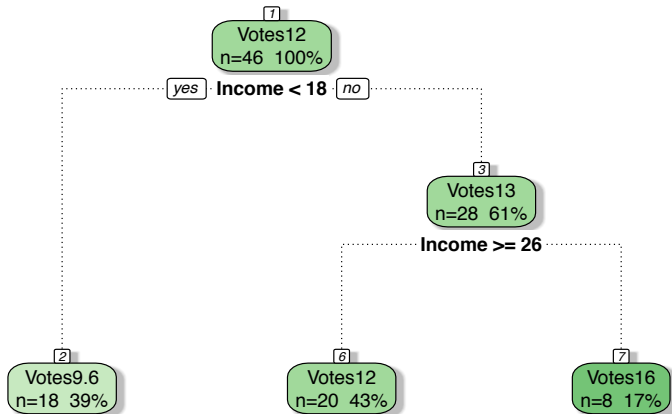
- Race Dependence.



Regression tree for Votes vs Race.

# Decision Trees contd....

- Dependence on Mean Household Income of each state.



Regression tree for Votes vs Mean Household Income.

# Inferences

US - Elections

Pramod  
Padmanabhan

Plan of the  
Talk

Introduction  
The Problem

Data and  
Variables

Visualizing the  
Dependencies

**Inferences**

Outlook

- Trump's national average is 12 % of the Republican votes.

# Inferences

US - Elections

Pramod  
Padmanabhan

Plan of the  
Talk

Introduction  
The Problem

Data and  
Variables

Visualizing the  
Dependencies

**Inferences**

Outlook

- Trump's national average is 12 % of the Republican votes.
- States with a significant Black and Hispanic population have voted below Trump's national average while he is way above average for the predominantly White states.

# Inferences

US - Elections

Pramod  
Padmanabhan

Plan of the  
Talk

Introduction  
The Problem

Data and  
Variables

Visualizing the  
Dependencies

Inferences

Outlook

- Trump's national average is 12 % of the Republican votes.
- States with a significant Black and Hispanic population have voted below Trump's national average while he is way above average for the predominantly White states.
- The upper middle class and the rich Republicans are the ones who endorse Trump more than the Republicans in the lower income groups.

# Inferences

US - Elections

Pramod  
Padmanabhan

Plan of the  
Talk

Introduction  
The Problem

Data and  
Variables

Visualizing the  
Dependencies

Inferences

Outlook

- Trump's national average is 12 % of the Republican votes.
- States with a significant Black and Hispanic population have voted below Trump's national average while he is way above average for the predominantly White states.
- The upper middle class and the rich Republicans are the ones who endorse Trump more than the Republicans in the lower income groups.
- States with more than half the counties urbanized support Trump more than the rural regions.



# Inferences contd....

US - Elections

Pramod  
Padmanabhan

Plan of the  
Talk

Introduction  
The Problem

Data and  
Variables

Visualizing the  
Dependencies

**Inferences**

Outlook

- Manufacturing and mining states have gone in favor of Trump compared to farming states and those which rely on State and Federal government jobs.

# Inferences contd....

US - Elections

Pramod  
Padmanabhan

Plan of the  
Talk

Introduction  
The Problem

Data and  
Variables

Visualizing the  
Dependencies

**Inferences**

Outlook

- Manufacturing and mining states have gone in favor of Trump compared to farming states and those which rely on State and Federal government jobs.
- States with a high percentage of unemployment support Trump more and the same goes for states with a low amount of education.

# Inferences contd....

US - Elections

Pramod  
Padmanabhan

Plan of the  
Talk

Introduction  
The Problem

Data and  
Variables

Visualizing the  
Dependencies

Inferences

Outlook

- Manufacturing and mining states have gone in favor of Trump compared to farming states and those which rely on State and Federal government jobs.
- States with a high percentage of unemployment support Trump more and the same goes for states with a low amount of education.
- Regions with a high presence of the KKK klans have voted with a percentage above Trump's national average. The surprising thing is this inference was made from Klan concentration data from the early part of the 20th century, when the Klans were most active, and they still seem to have an effect on present day voters.

# Outlook

US - Elections

Pramod  
Padmanabhan

Plan of the  
Talk

Introduction  
The Problem

Data and  
Variables

Visualizing the  
Dependencies

Inferences

Outlook

- This analysis can be performed for Hillary Clinton as well. This would give us a way to make a comparison between the two candidates which gives a way to predict the outcome of the elections.

US - Elections

Pramod  
Padmanabhan

Plan of the  
Talk

Introduction  
The Problem

Data and  
Variables

Visualizing the  
Dependencies

Inferences

Outlook

**Thank you !**