Agenda - Today

- Correspondence anlaysis
- Max-diff
- Driver analysis
- Segmentation

Software

The start-of-the-art data science app.

The world's largest open source data science/stats package.

Difficult-to-use.

Not designed by market researchers/marketing analytics people.

Designed to make it easier to extend R (i.e., contribute to the open source project)

Designed to make it easy to use R for data science and building online dashboards.

Designed to make it easy to use R for market research.

Agenda – This session

- What is correspondence analysis?
- When to use it
- How it works
- Correct interpretation
- Analyzing trend data

Trump approval, voting, and demographics

Q1. Are you …

	%
Male	49%
Female	51%

Q2 How old are you?

	%
Less than 18 years	0%
18 to 24 years	11%
25 to 34 years	18%
35 to 44 years	18%
45 to 54 years	19%
55 to 64 years	17%
65 years or over	17%

Q10 Did you…

	%
Vote for Clinton	43%
Vote for Trump	31%
Vote for another candidate	10%
Not vote	11%
Don't know	0%
I refuse to answer this question	5%

Q13 Do you approve or disapprove of the way Donald Trump is handling his job as president?

|	%
Approve	30%
Disapprove	53%
No Opinion	17%

Q11 In 2012, you remember that Obama ran for President on the Democratic ticket against Romney for the Republicans. Did you…

|	%
Vote for Obama	47%
Vote for Romney	28%
Vote for another candidate	5%
Not vote	15%
Don't know	1%
I refuse to answer this question	4%

How to read a correspondence analysis map

- Compare between row labels based on distances (if *row principal* or *principal normalization*).

- Compare columns labels based on distances (if *column principal or principal normalization).*

- To compare a row label to a column label:
  - Look at the length of the line connecting the row label to the origin. Longer lines indicate that the row label is highly associated with some of the column labels (i.e., it has at least one high residual).
  - Look at the length of the label connecting the column label to the origin. Longer lines again indicate a high association between the column label and one or more row labels.
  - Look at the angle formed between these two lines. Very small angles indicate association. 90 degree angles indicate no relationship. Angles near 180 degrees indicate negative associations (if *row principal, column principal,* or *symmetrical (1/2)*).

- Always check conclusions using the raw data

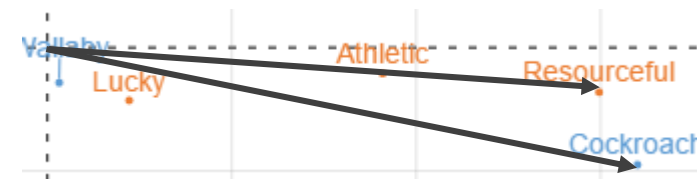- The lower the variance explained, the more we need to check the raw data
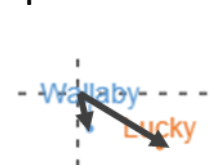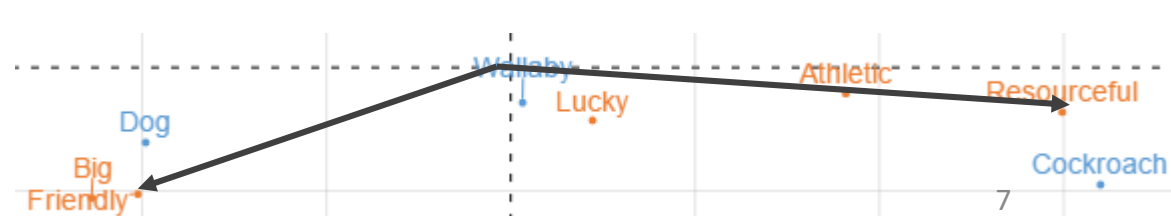
Strong positive association
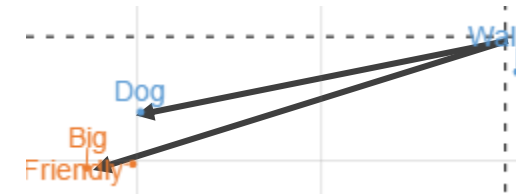
No association
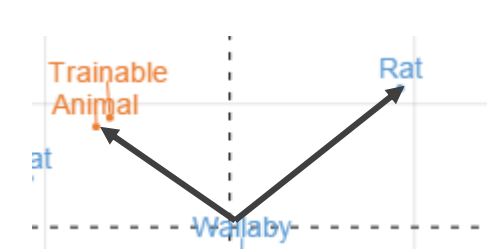
Strong positive association

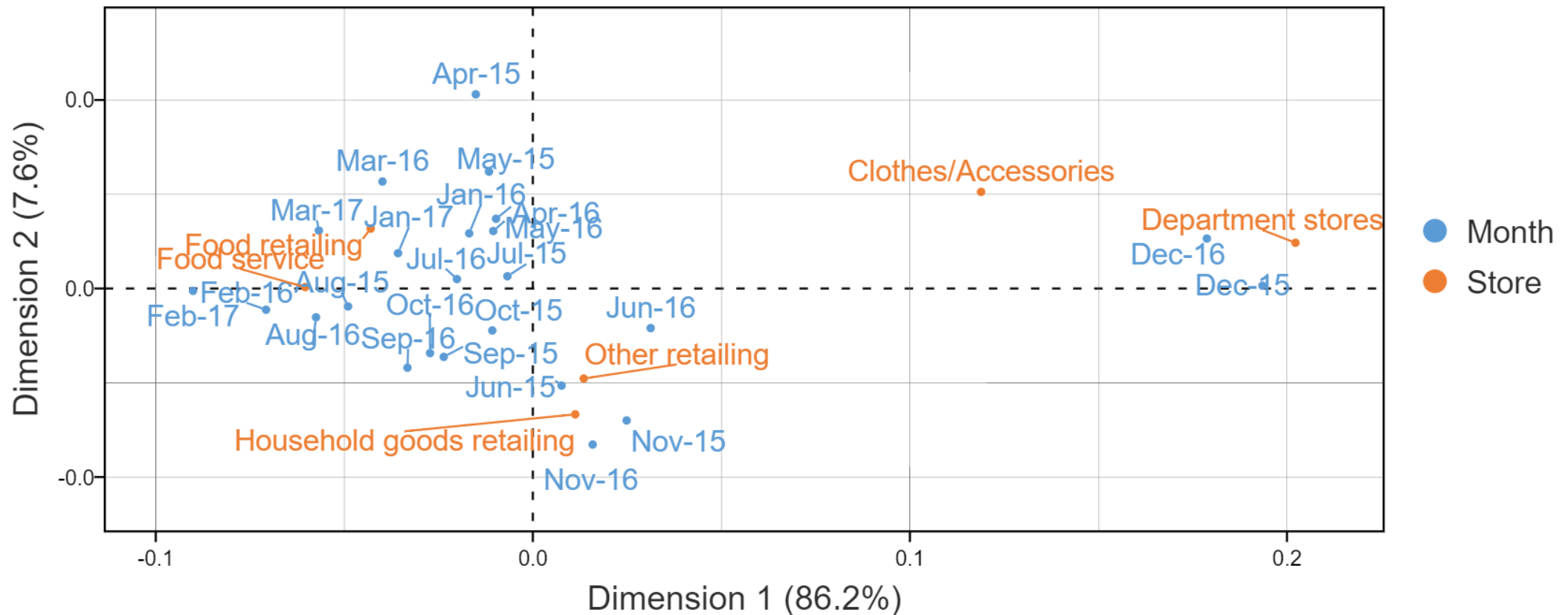Very weak positive association

Strong negative association

7

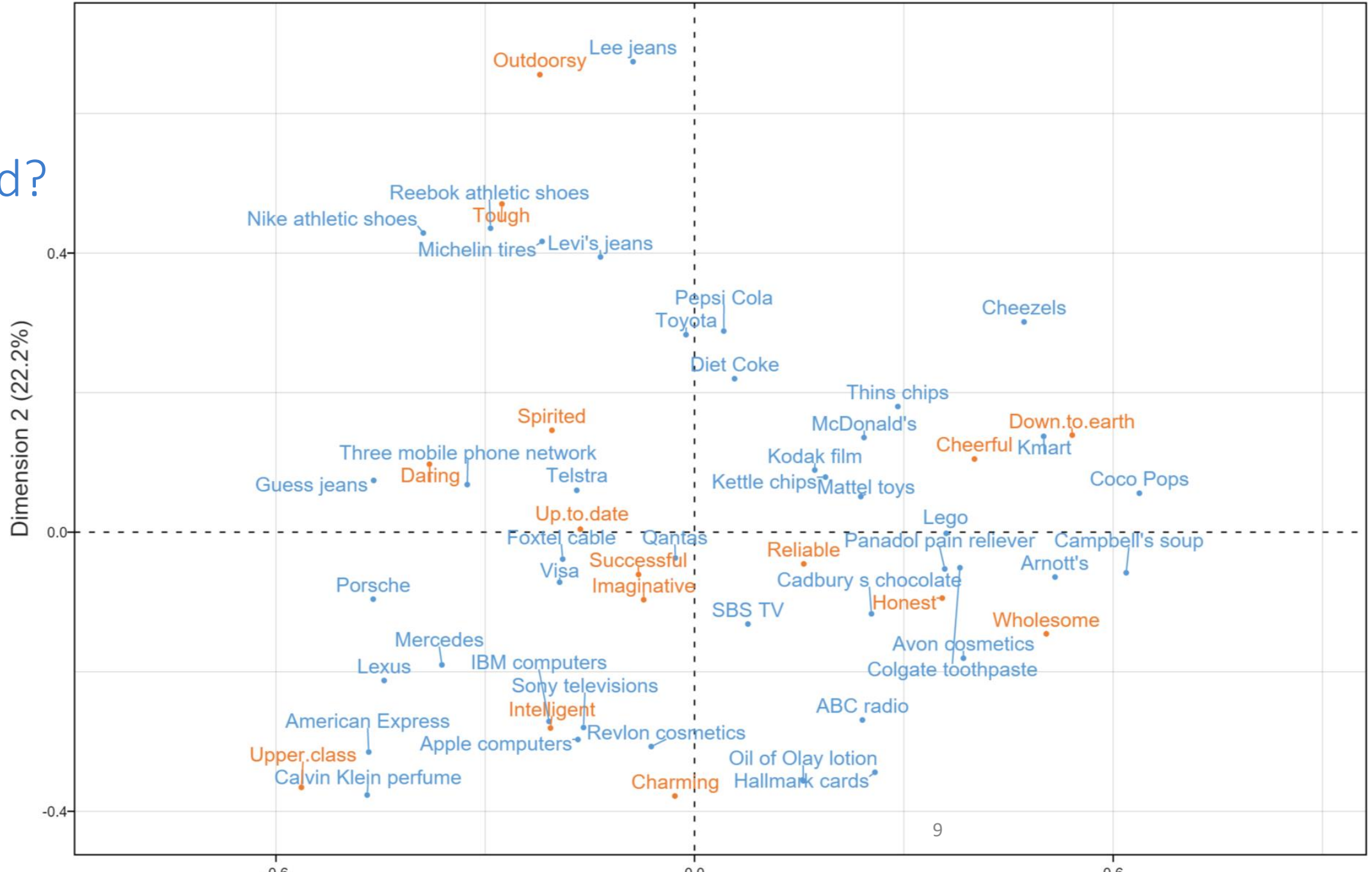# What can we infer about department store sales?

Correspondence analysis of sales data over time

How is Qantas positioned?

A perceptual map with Dimension 1 on the x-axis (ranging from -0.6 to 0.6) and Dimension 2 (22.2%) on the y-axis (ranging from -0.4 to above 0.4). Brands and attributes are plotted:

Outdoorsy, Lee jeans, Reebok athletic shoes, Tough, Nike athletic shoes, Michelin tires, Levi's jeans, Pepsi Cola, Toyota, Cheezels, Diet Coke, Thins chips, Spirited, McDonald's, Down.to.earth, Three mobile phone network, Cheerful Kmart, Kodak film, Daring, Telstra, Kettle chips, Mattel toys, Coco Pops, Guess jeans, Up.to.date, Lego, Foxtel cable, Qantas, Reliable, Panadol pain reliever, Campbell's soup, Successful, Visa, Cadbury s chocolate, Arnott's, Imaginative, Honest, Porsche, SBS TV, Wholesome, Mercedes, Avon cosmetics, Lexus, IBM computers, Colgate toothpaste, Sony televisions, ABC radio, Intelligent, American Express, Revlon cosmetics, Apple computers, Upper.class, Oil of Olay lotion, Calvin Klein perfume, Charming, Hallmark cards

9

What is the topline conclusion here?

Multiple correspondence analysis of Gender, Age, 2016, 2012, Trump
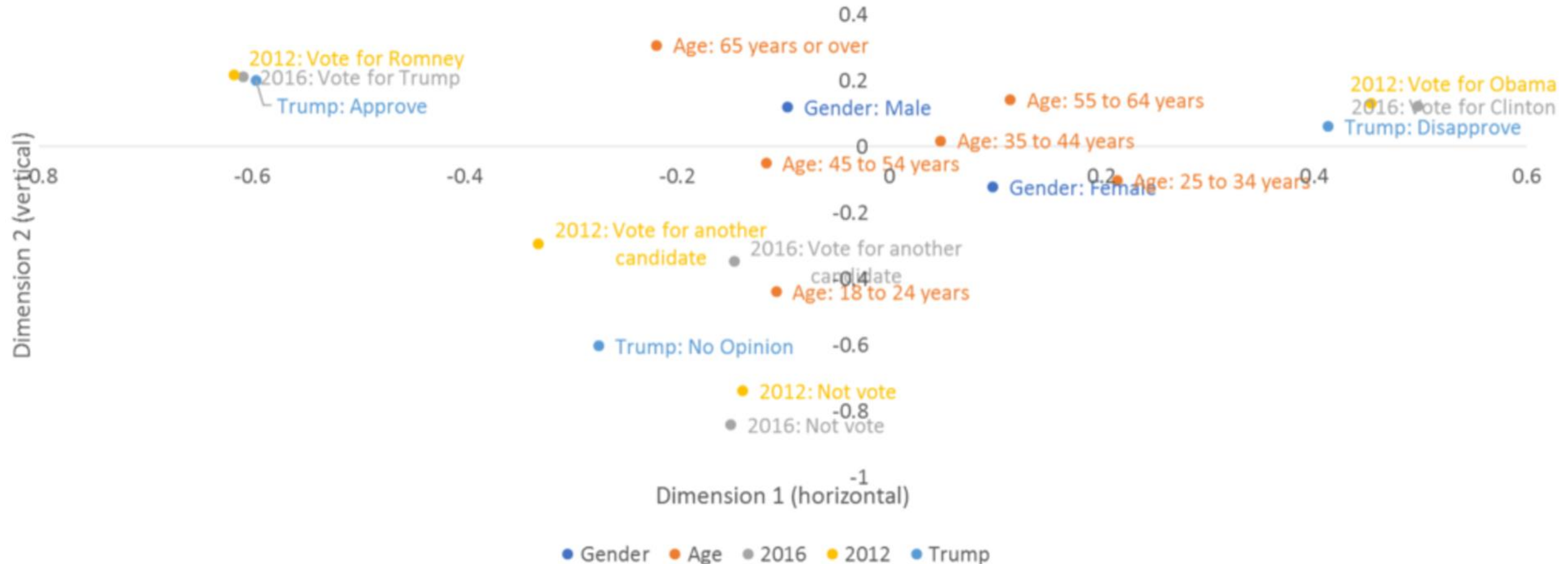
Tell me about people that did not vote?

Multiple correspondence analysis