

Project 1: UMD

Yangjianchen Xu

Overview

- Background
- Purpose
- Related questions
- Data processing
- Answers for questions
- Conclusions
- Future work

Background

- The programs of **Urban Ministries of Durham (UMD)** end homelessness by providing neighbors with emergency shelter and case management to help them overcome barriers such as unemployment, medical and mental health problems, past criminal convictions and addiction.
- The data is provided by UMD, including 79838 observations with 18 variables such as **Date**, **Client File Number**, **Food Pounds** and **Clothing Items**.

Purpose

- Overall purpose: helping UMD determining the amount of food and clothing items they need to provide for the coming years.

Related questions

- What is the relationship among Food Pounds, Clothing Items and Number of people in the family for which food was provided?
- How does time influence the amount of food and clothing items?
- What is the amount of food and clothing items need to be provided in 2019?

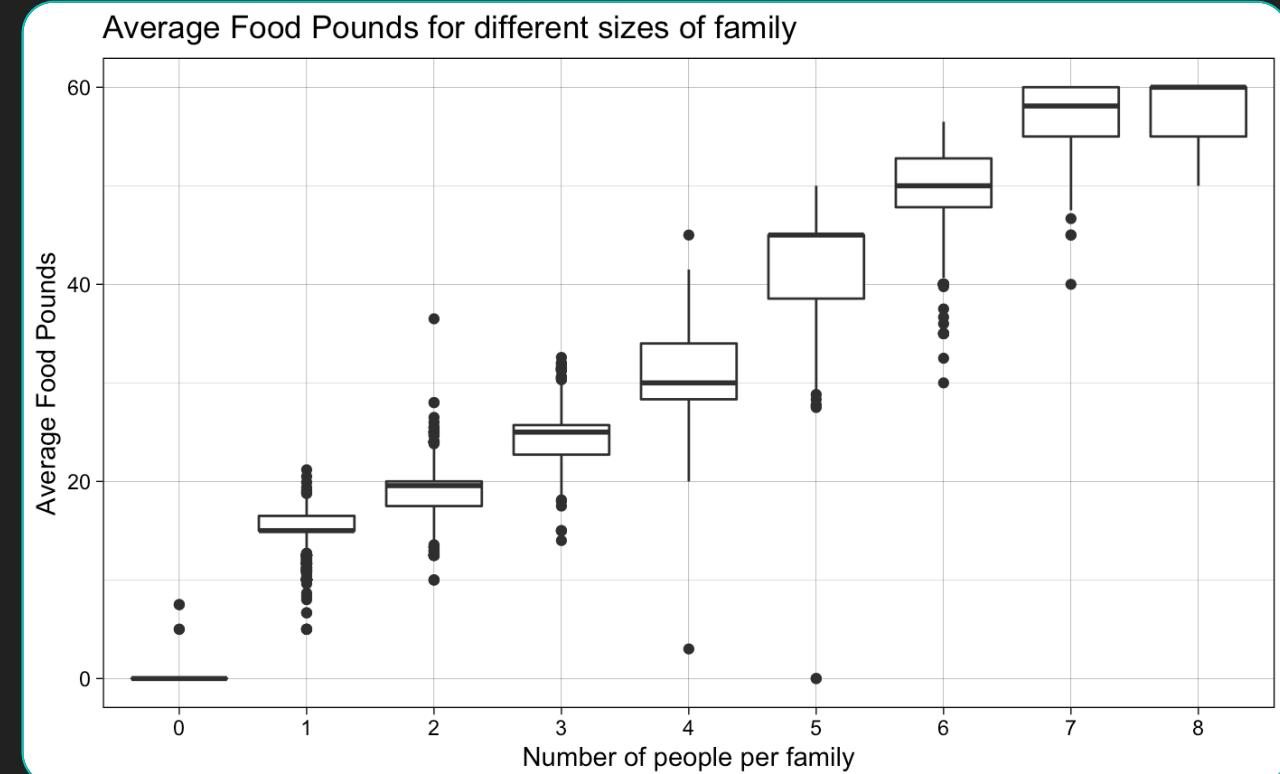
Data processing

- Select variables: Date, Client File Number, Food Pounds, Clothing Items and Number of people in the family for which food was provided (Number of people per family).
- Discard observations with NA in these variables.
- Discard outliers.

▲	Date	CFN	food	clothing	number
1	1/22/2009	212	20	5	3
2	1/29/2009	738	25	26	4
3	1/20/2009	3455	40	39	6
4	2/19/2008	1814	0	0	0
5	5/1/1931	4	15	10	1

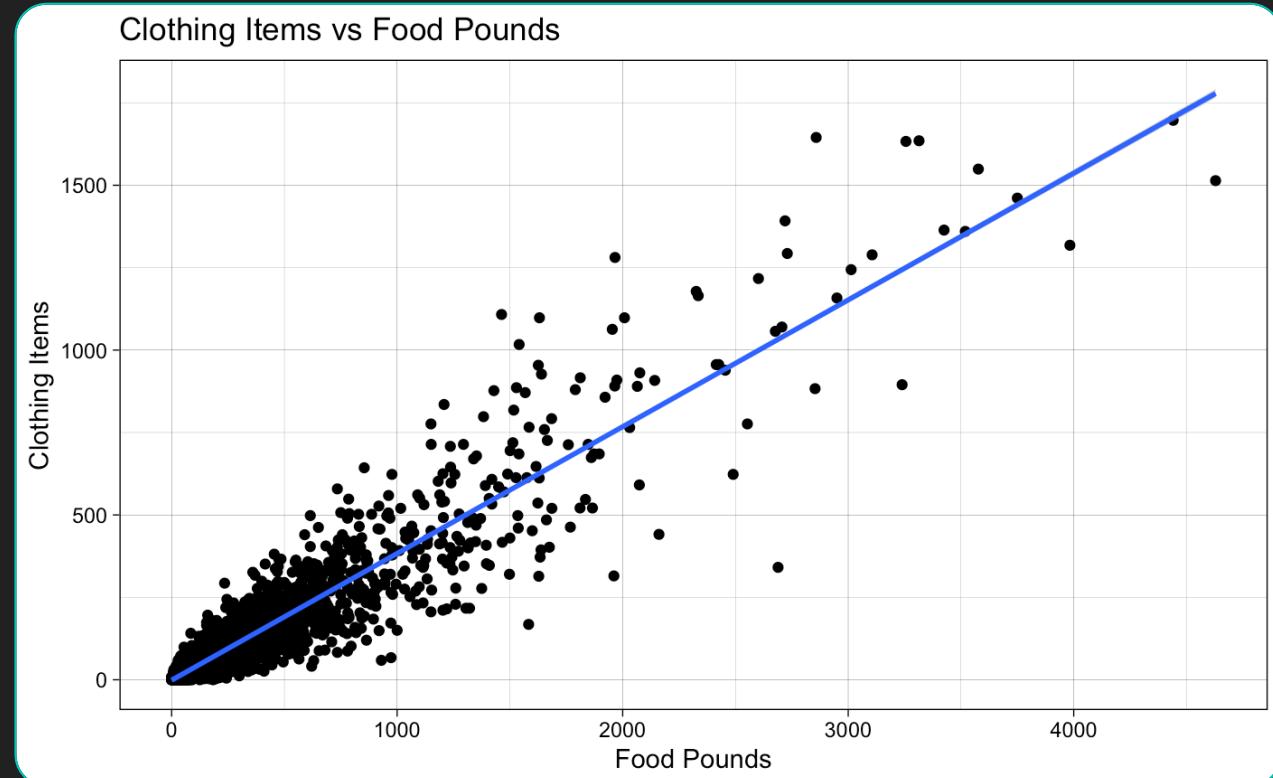
Question 1

- What is the relationship among Food Pounds, Clothing Items and Number of people in the family for which food was provided?
- Average Food Pounds: the average food amount received per family.
- There is a positive correlation.



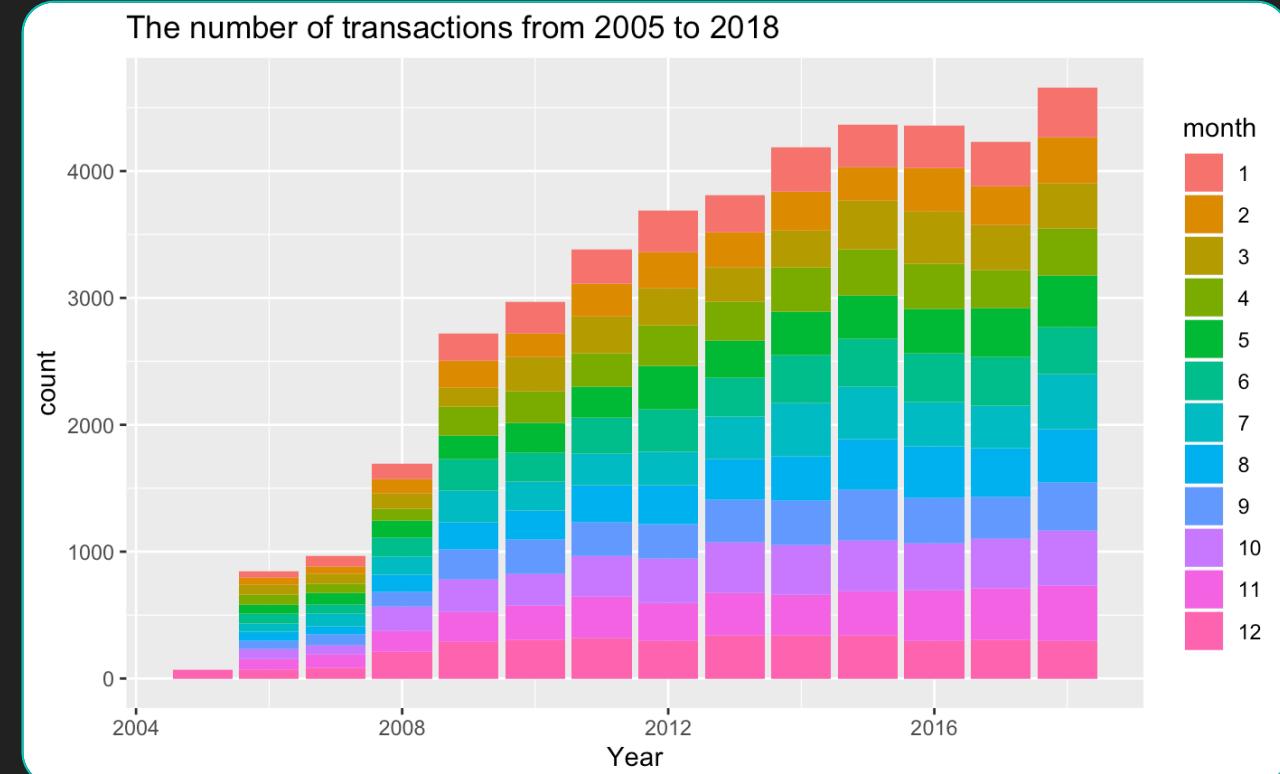
Question 1

- Clothing Items: the number of clothing items per transaction
- Food Pounds: the amount of food per transaction
- The variance become larger when Food Pounds grows.
- Data transformation is required.

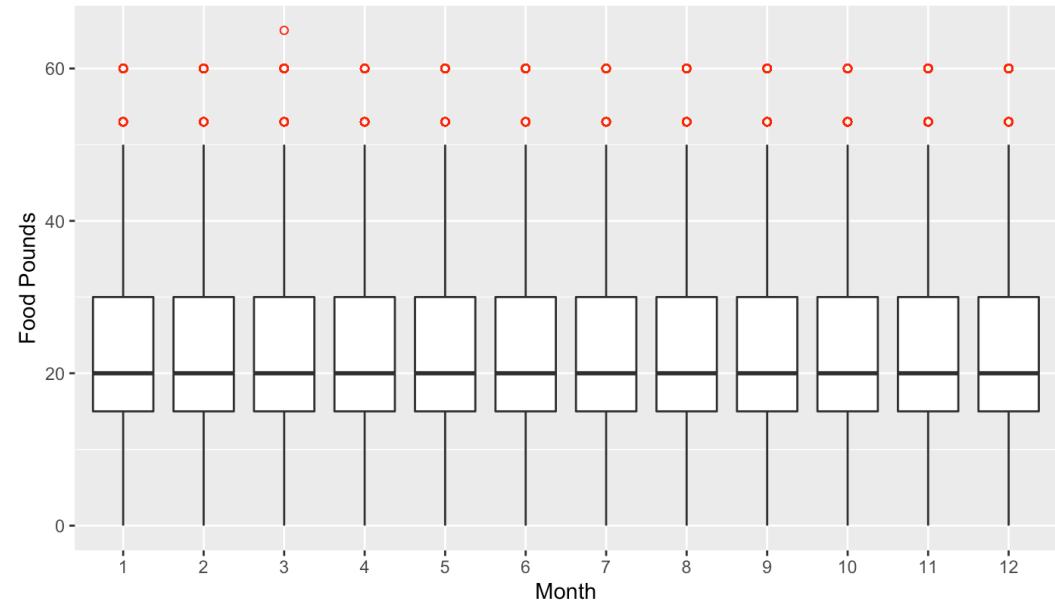


Question 2

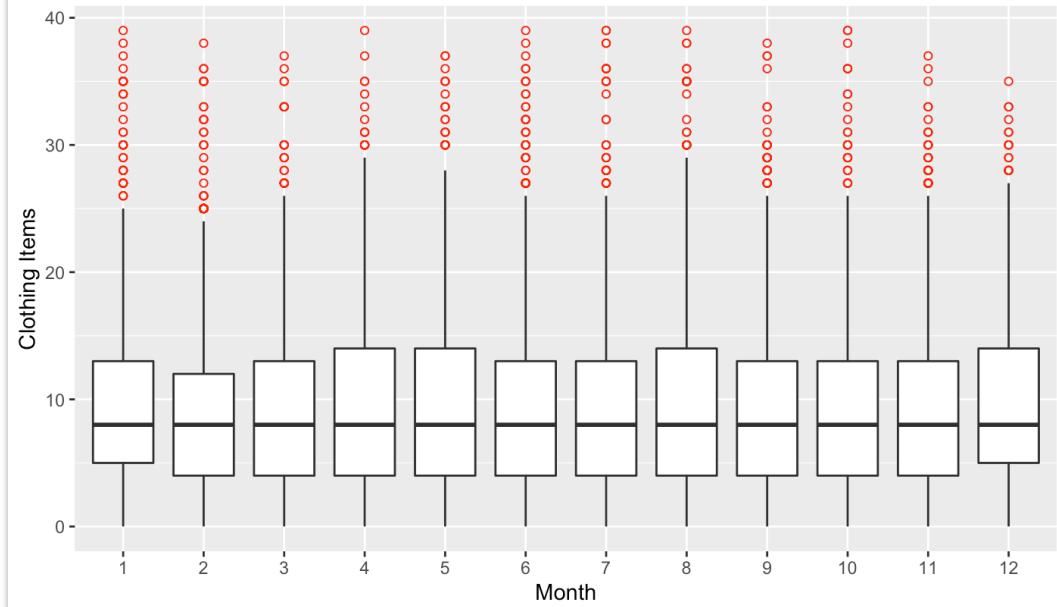
- How does time influence the amount of food and clothing items?
 - The number of transactions is increasing by year.
 - In each year, the number of transactions seems to be monthly equal.



Amount of food for different months



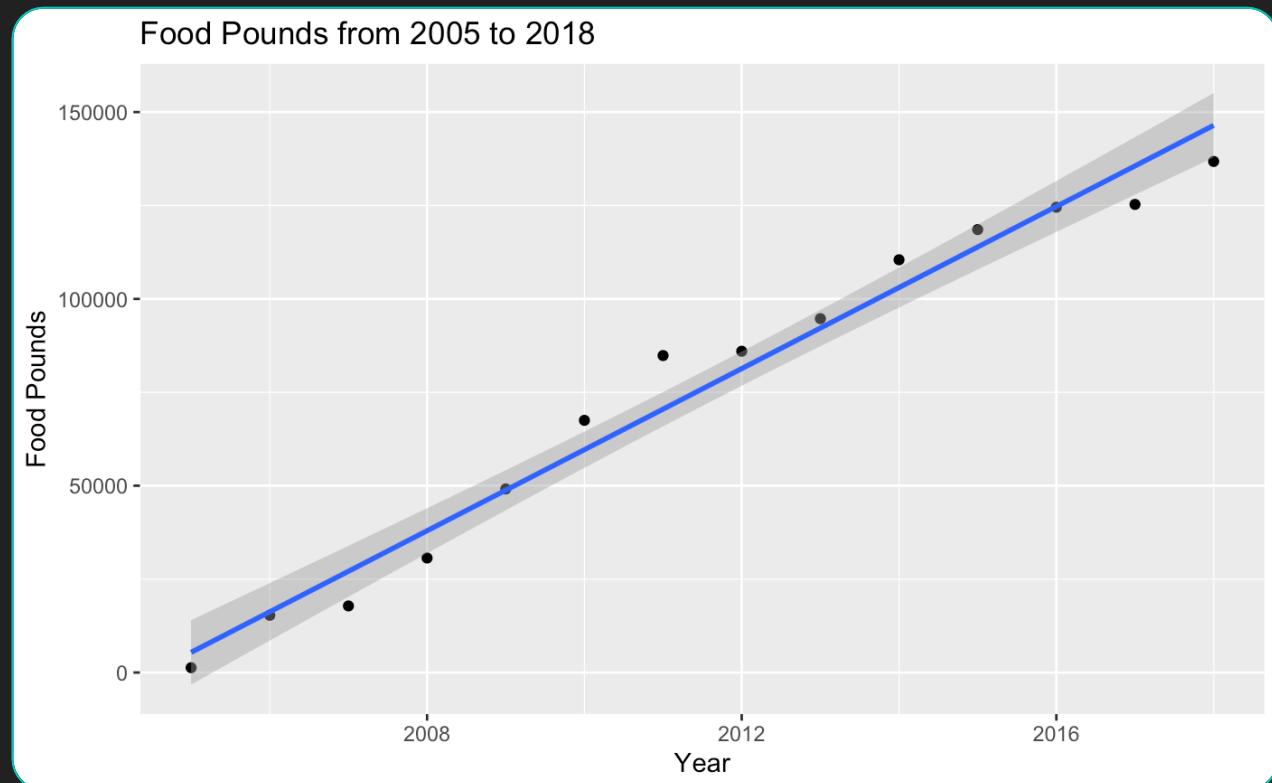
Number of clothing items for different months



Distribution of food and clothing items

Question 2

- The amount of food and year seem to have a linear correlation.
- Use this linear model to predict the amount of food need to be provided for the coming years.

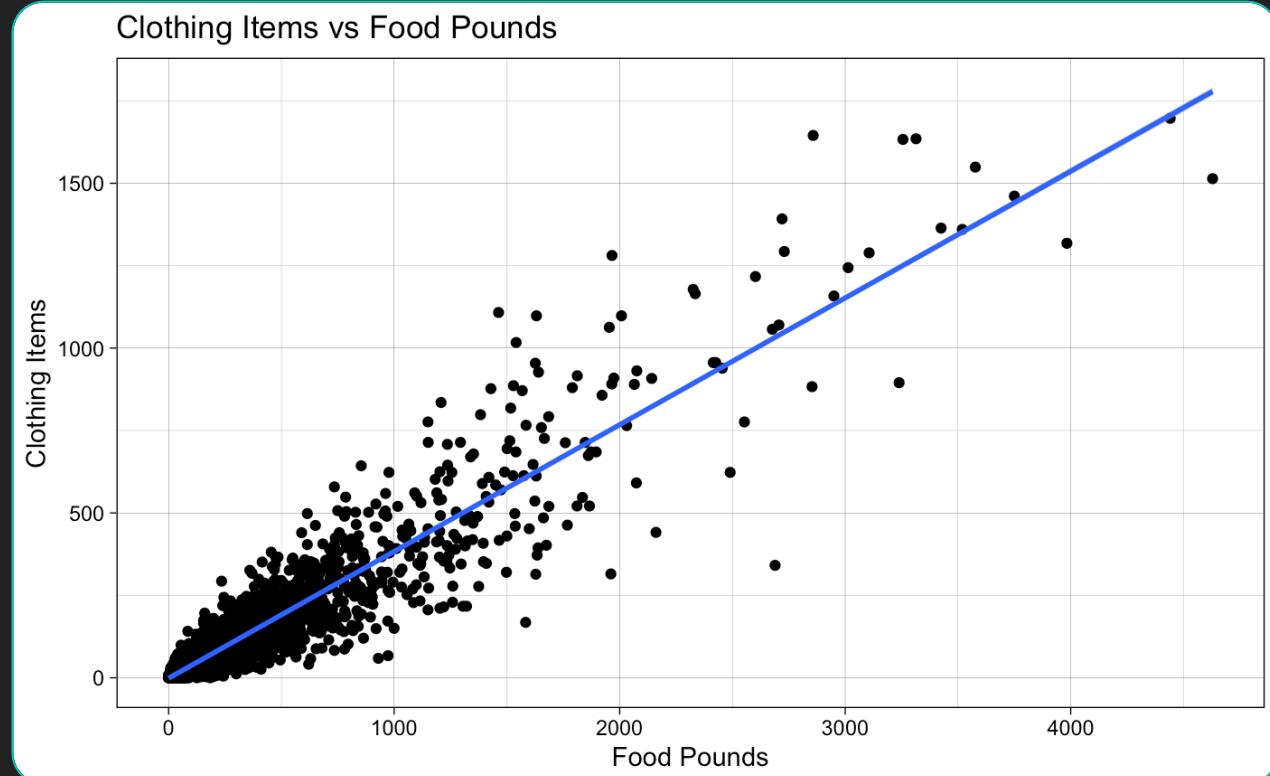


Question 3

- Fitted model: $Food = 10849 * Year - 21747332$
- For 2019: $Food(\widehat{2019}) = 156799$

Question 3

- How about the number of clothing items in 2019?
- Use weighted linear model.



Conclusion

- Main conclusions:
 - The amount of food and year have a linear correlation: $Food = 10849 * Year - 21747332$
 - The amount of food for the coming years can be predicted by this model.
 - It takes further work to predict the number of clothing item.
- Secondary conclusions:
 - There is a positive correlation between Average Food Pounds and the size of family.
 - The number of transactions is increasing by year.
 - In each year, the number of transactions are roughly the same every month.
 - The distribution of food and clothing items are roughly the same for different months.

Future work

- Predict the number of clothing item need to be provided for the coming years by analyzing the relationship between Food Pounds and Clothing Items.