

# Data Journalism

METHODS AND CASE STUDIES FROM MONGOLIA

# About Robert Ritz



Director of LETU  
Mongolia



Data Scientist at  
Mongol.ai



Creator of Mongolian  
Data Stories



Lived in Mongolia  
since 2013



# Mongolian Data Stories

Using data to tell Mongolia's story

[MACHINE LEARNING](#)

[DATA VISUALIZATION](#)

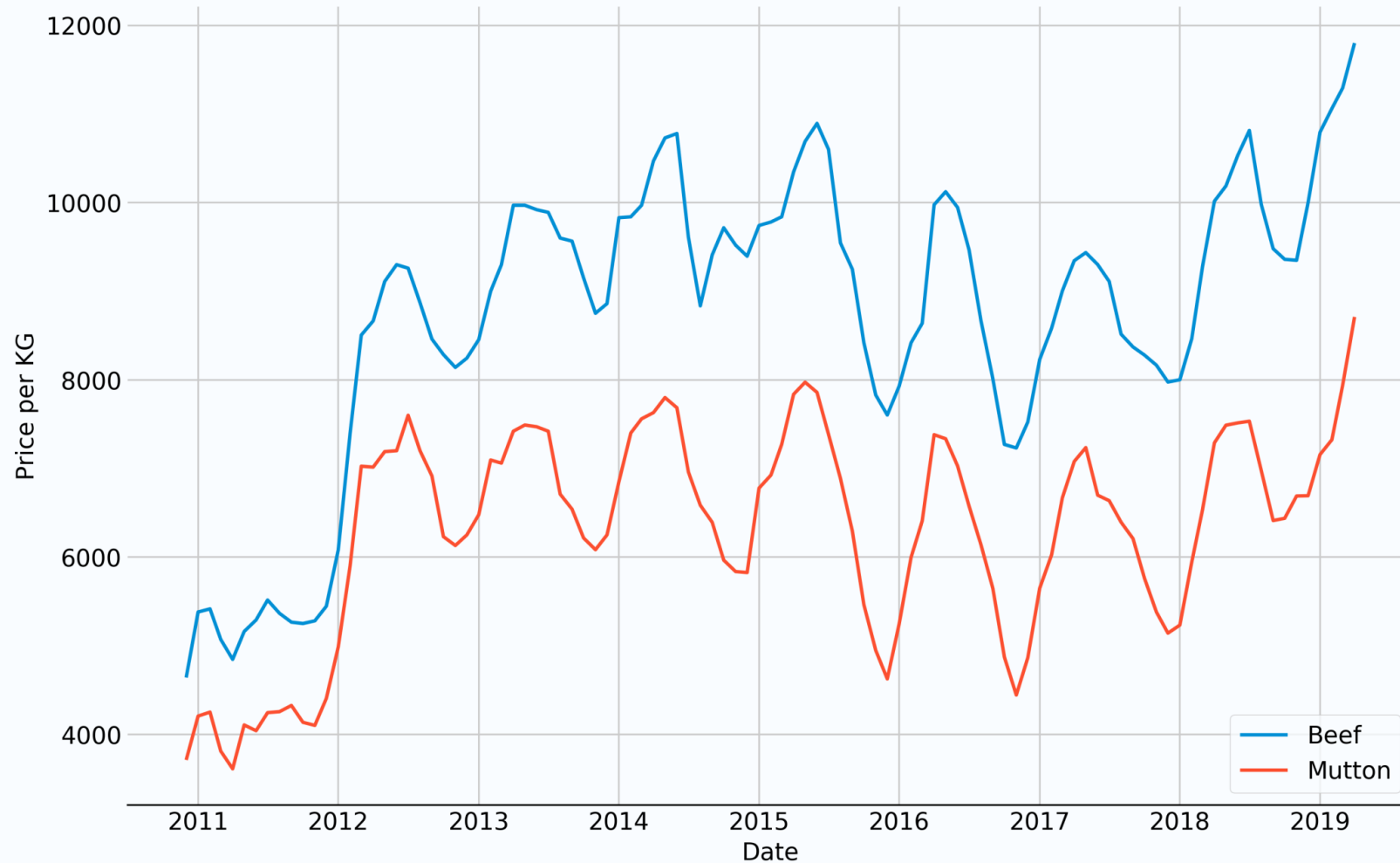
[ARCHIVE](#)

[МОНГОЛ ХЭЛ](#)

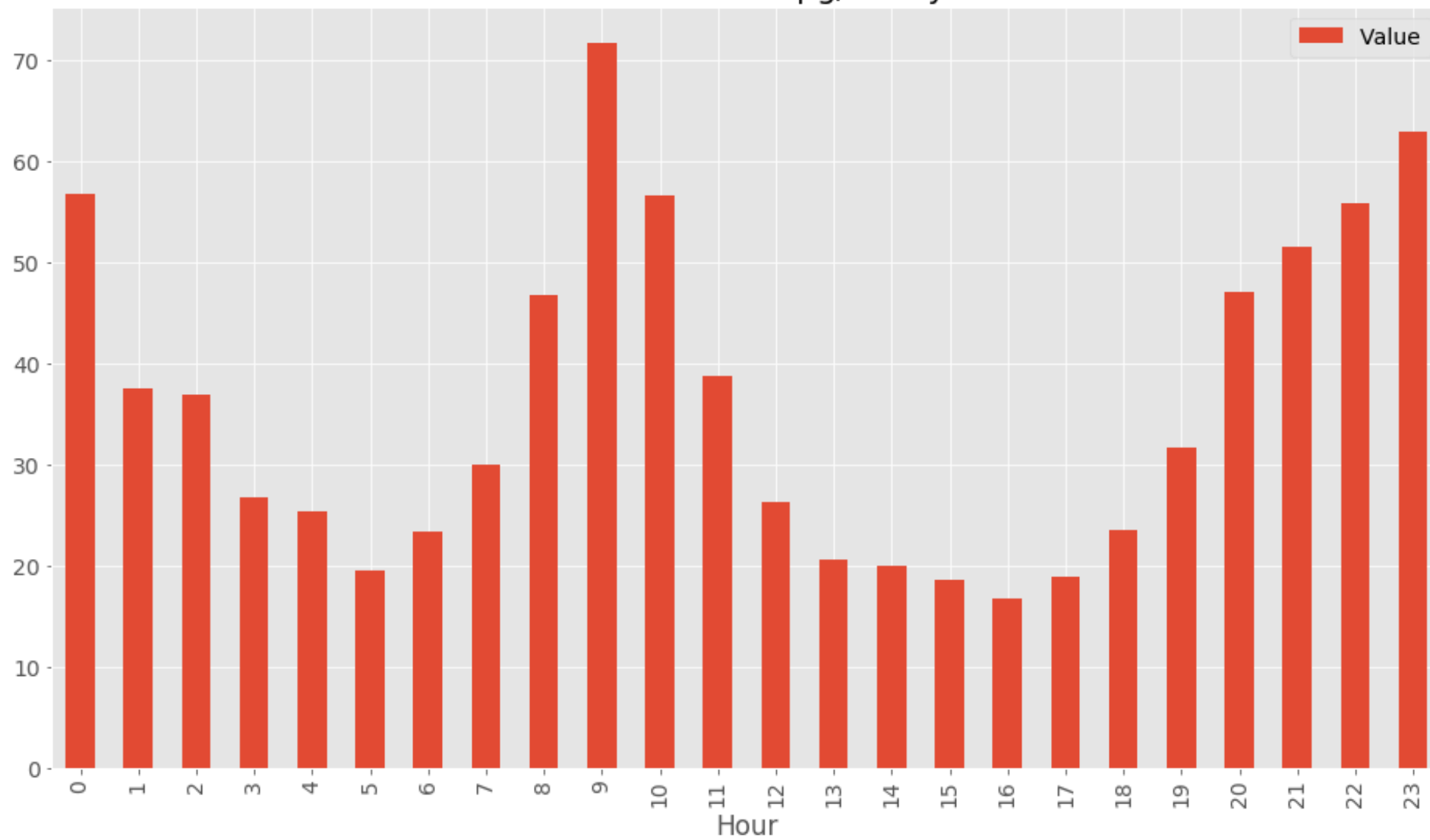
[ABOUT MDS](#)

# Roller Coaster of Meat Prices

Dec 2011 to Apr 2019



Mean PM2.5 Level in  $\mu\text{g}/\text{m}^3$  by Hour



# Goals



Understand the data analysis process



Understand ways to create reproducible reporting

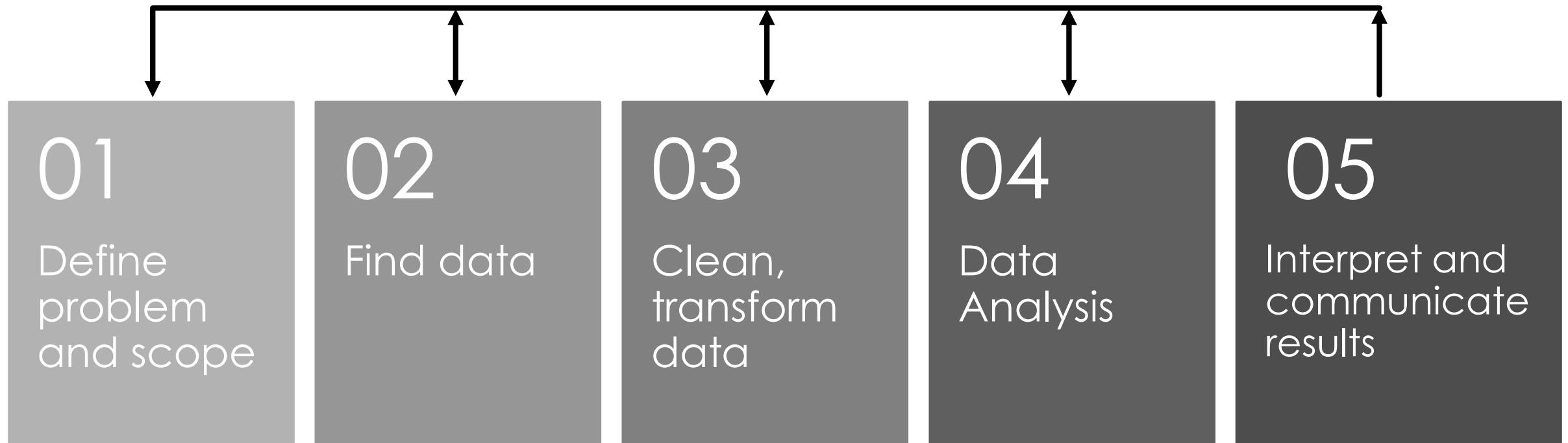


See how this works through case studies

# Kahoot!

Data Journalism Quiz!

# Data Analysis Process





# Case Studies

DATA ANALYSIS PROCESS  
REPRODUCIBILITY

Problem:  
Perception of  
Chronic Youth  
Unemployment

Redefined  
problem:

How is unemployment  
distributed through the  
population?

Age group	Sex	2009	2010	2011	2012	2013
Total	Total	11.6	9.9	7.7	8.2	7.9
	Male	11.6	10.5	8.1	8.4	7.6
	Female	11.6	9.2	7.4	8.1	8.3
15-19	Total	22.8	18	14	8	13.6
	Male	23.1	17.5	14.6	6.9	11.1
	Female	22.4	18.7	13	9.3	17.6
	Total	21.7	20	16.1	14.1	13.1

# Data Collection

1212.MN (NATIONAL STATISTICAL OFFICE)

UNEMPLOYMENT RATE, BY SEX, AGE GROUP

# Clean, Transform Data



Is this data ready to be analyzed?



Often statistical data is in “short” form.



Most data visualization software needs “long” form data.



# Analyze Data

No additional analysis needed

May want context to understand the data

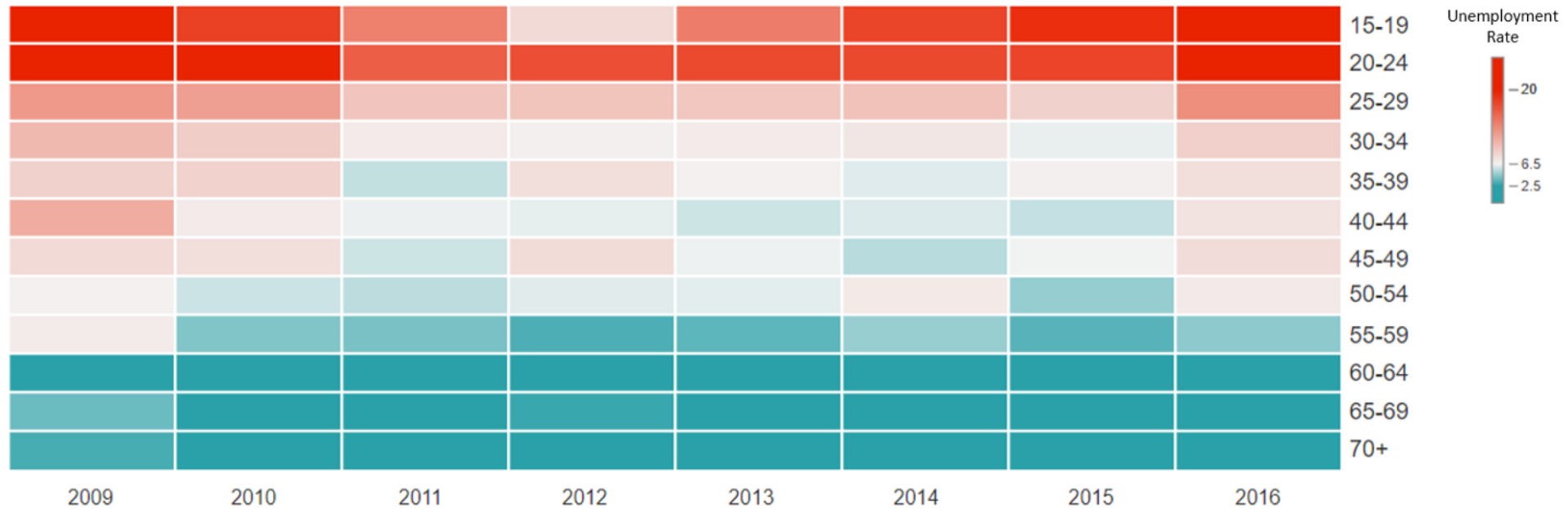
- How is unemployment rate calculated?

# Unemployment in Mongolia

## By Age Group

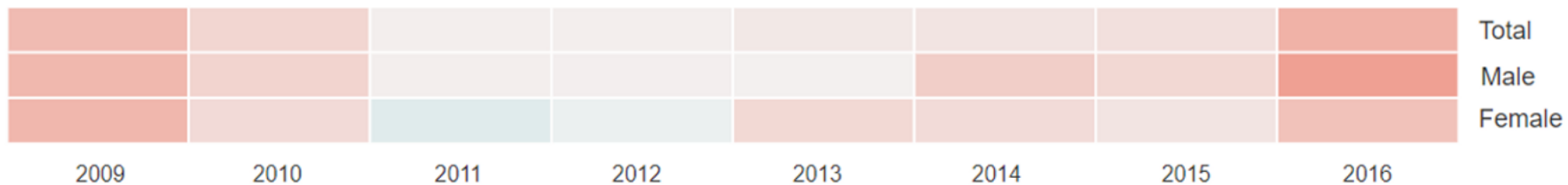
Aside from a low in 2012, youth unemployment remains very high.

How much of the state budget is allocated to old vs young?



## By Gender

While males have a slightly higher unemployment rate, the disparity is much higher by age group.





## Tools used:

- ▶ Excel – data cleaning and analysis
- ▶ Data Illustrator – data communication



Mongolia is a  
peaceful country

Redefined  
problem:

IS THE MURDER RATE  
HIGHER OR LOWER  
THAN THE USA?



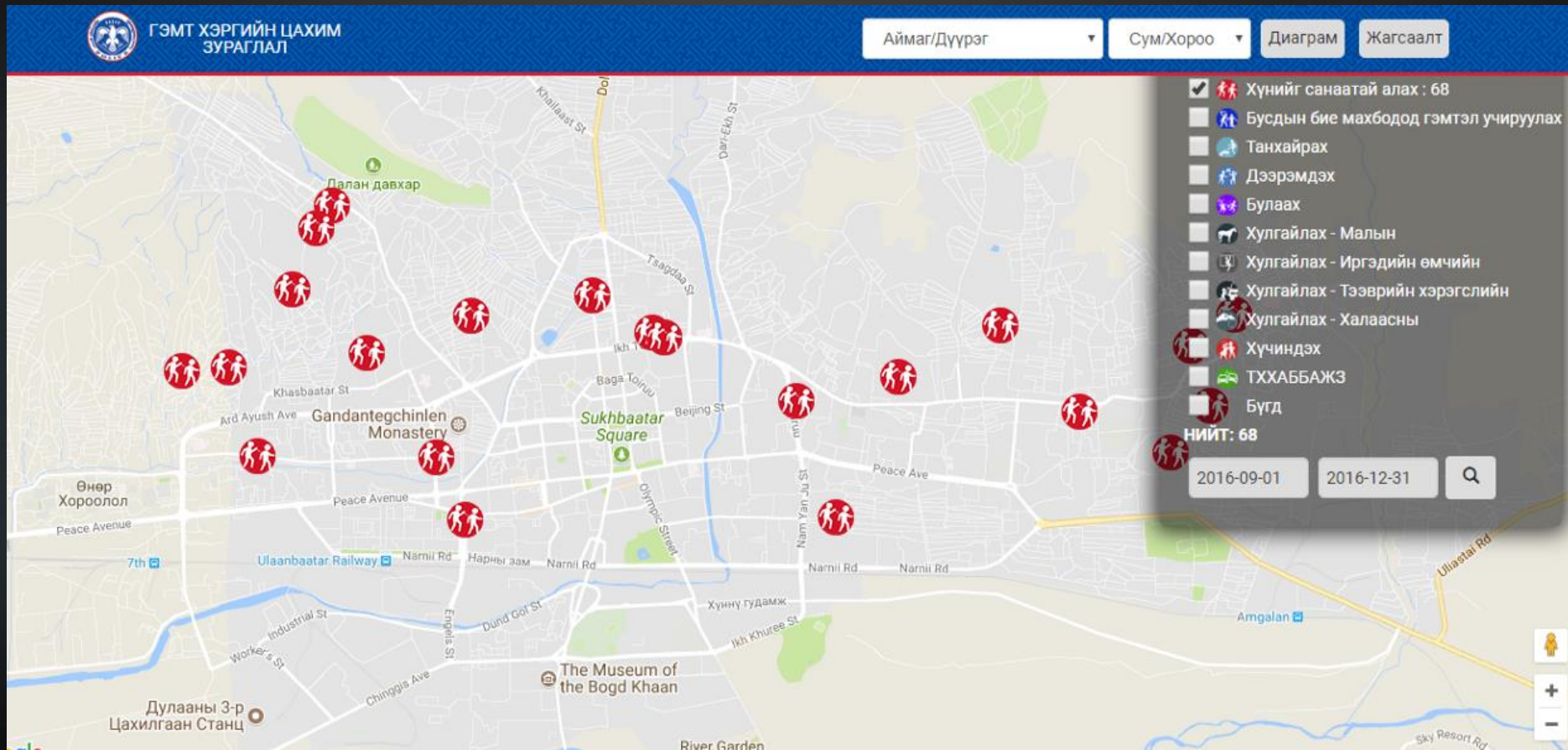
Poll:

In 2016 was the Mongolian murder rate higher than the United States murder rate?



# Data:

- ▶ Official crime statistics are released the year after
- ▶ Alternative source: <http://crimemap.police.gov.mn>



# Cleaning and Analysis



Data is from every aimag (province) of Mongolia



Dataset is over 56,000 rows, messy



Shows every crime report in 2015, 2016



How do we show data of this type?

6.48

2015 Intentional Homicide

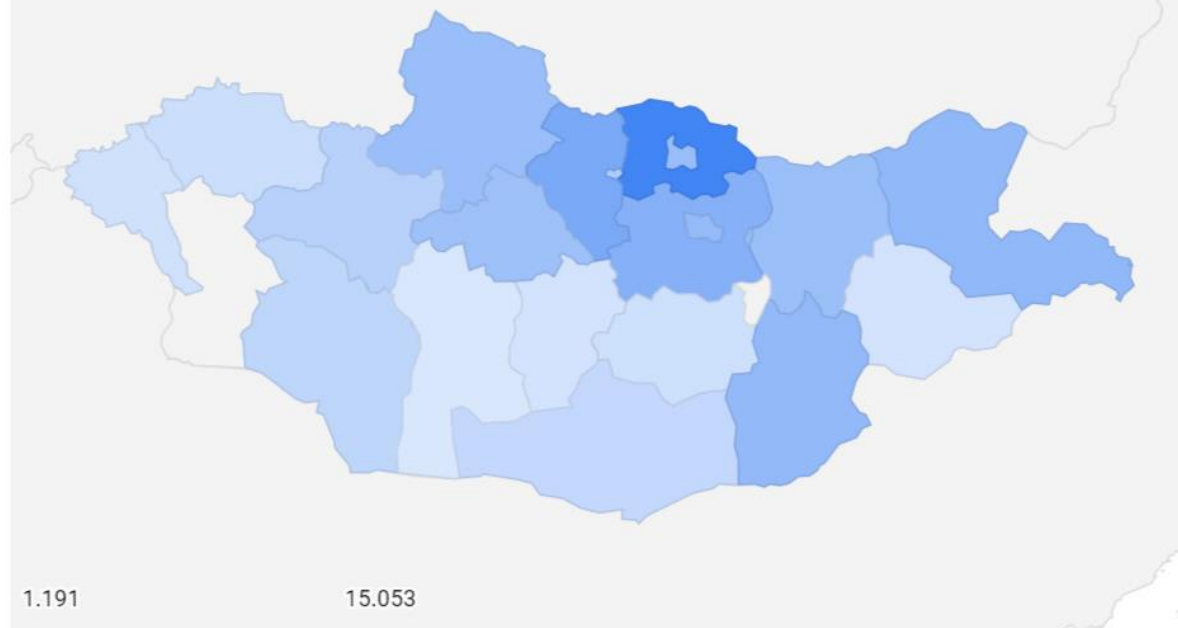
## Murder Rate by Aimag

Murders per 100,000 inhabitants

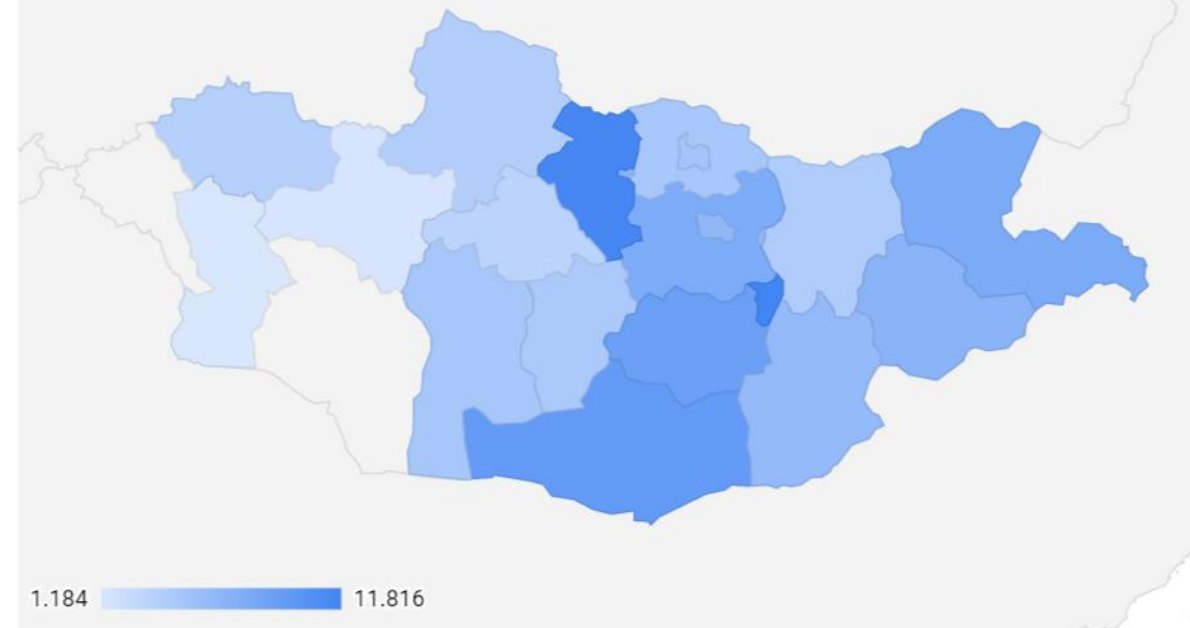
5.72

2016 Intentional Homicide

2015 Murder Rate by Aimag



2016 Murder Rate by Aimag



	Хэргийн дугаар ▾	Хэргийн огноо	Хот/Аймаг	Дүүрэг	Хороо/Сум
1.	2015260001491	Mar 3, 2015	Улаанбаатар	Сонгинохайрхан д...	20-р хороо
2.	2015260001228	Feb 25, 2015	Улаанбаатар	Хан-Уул	10-р хороо
3.	2015260001002	Feb 13, 2015	Улаанбаатар	Сонгинохайрхан д...	22-р хороо
4.	2015260000440	Jan 19, 2015	Улаанбаатар	Сонгинохайрхан д...	19-р хороо

1 - 10 / 193 &lt; &gt;

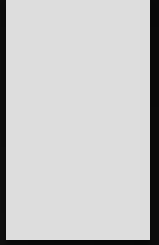
	Хэргийн дугаар ▾	Хэргийн огноо	Хот/Аймаг	Дүүрэг	Хороо/Сум
1.	201726010003	Dec 31, 2016	Улаанбаатар	Сонгинохайрхан д...	19-р хороо
2.	201726010002	Dec 29, 2016	Улаанбаатар	Сонгинохайрхан д...	3-р хороо
3.	201722000119	Oct 30, 2016	Хөвсгөл	null	Бүрэнтогтох сум
4.	201706000001	Dec 30, 2016	Булган	null	Хангал сум

1 - 10 / 173 &lt; &gt;

# Result:

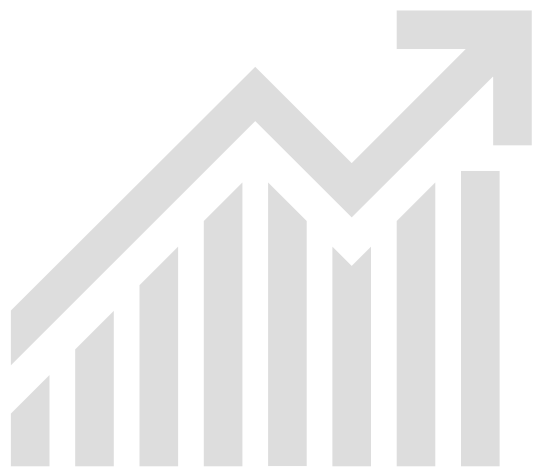
- ▶ 2016 Mongolia murder rate\*: 5.72 (5.66 according to UN-CTS)
- ▶ 2016 US murder rate: 5.35
  - ▶ UN-CTS reported 171 murders in 2016. Crimemap reported 173.
- ▶ 2016 data became available in 2018.
- ▶ Limited in what we can analyze from this data.
  - ▶ Lack of detailed information about crime.
  - ▶ Who was the victim? Theft has four categories, assault has one.
  - ▶ Was it closed or not? Did it lead to a prosecution?
- ▶ Website no longer works





## Tools used:

- ▶ Excel – data cleaning and analysis
- ▶ Google Data Studio



Is debt  
becoming  
unsustainable  
in Mongolia?

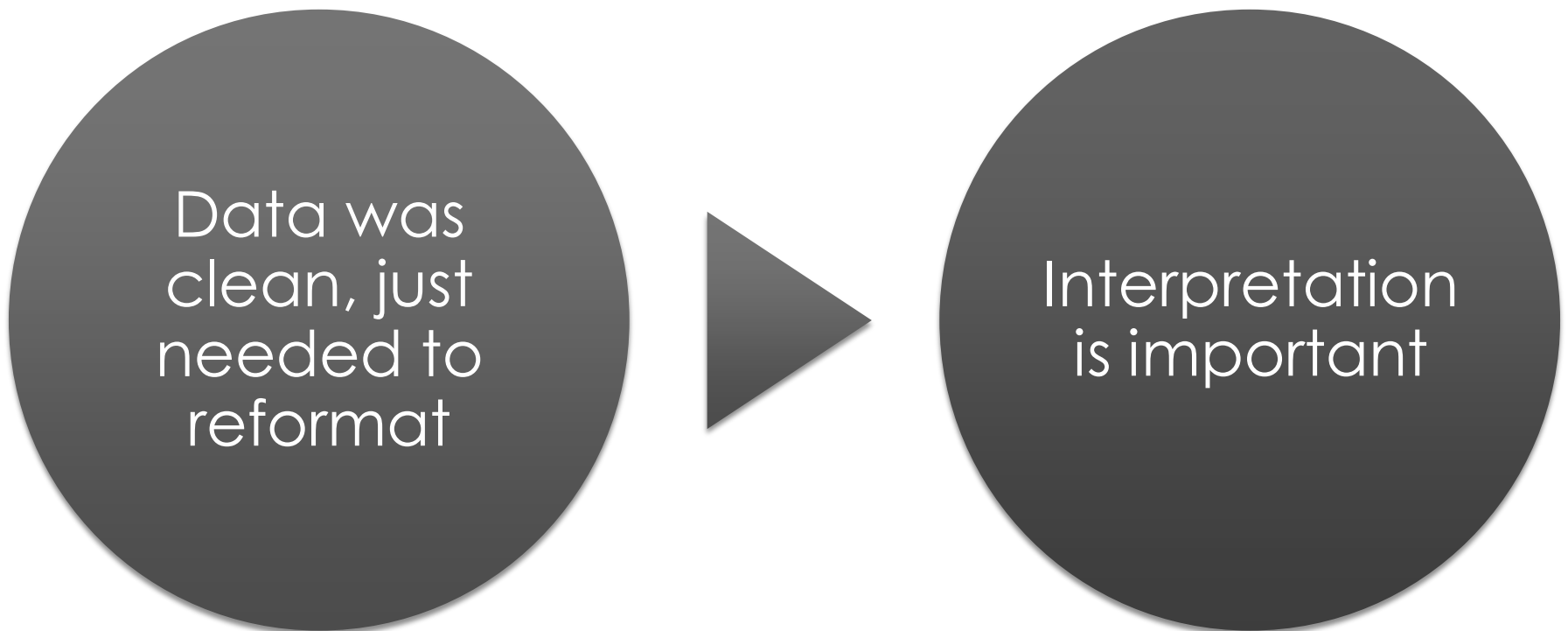
# Data



Outstanding loans and  
outstanding deposits



1212.mn (National Statistical  
Office)



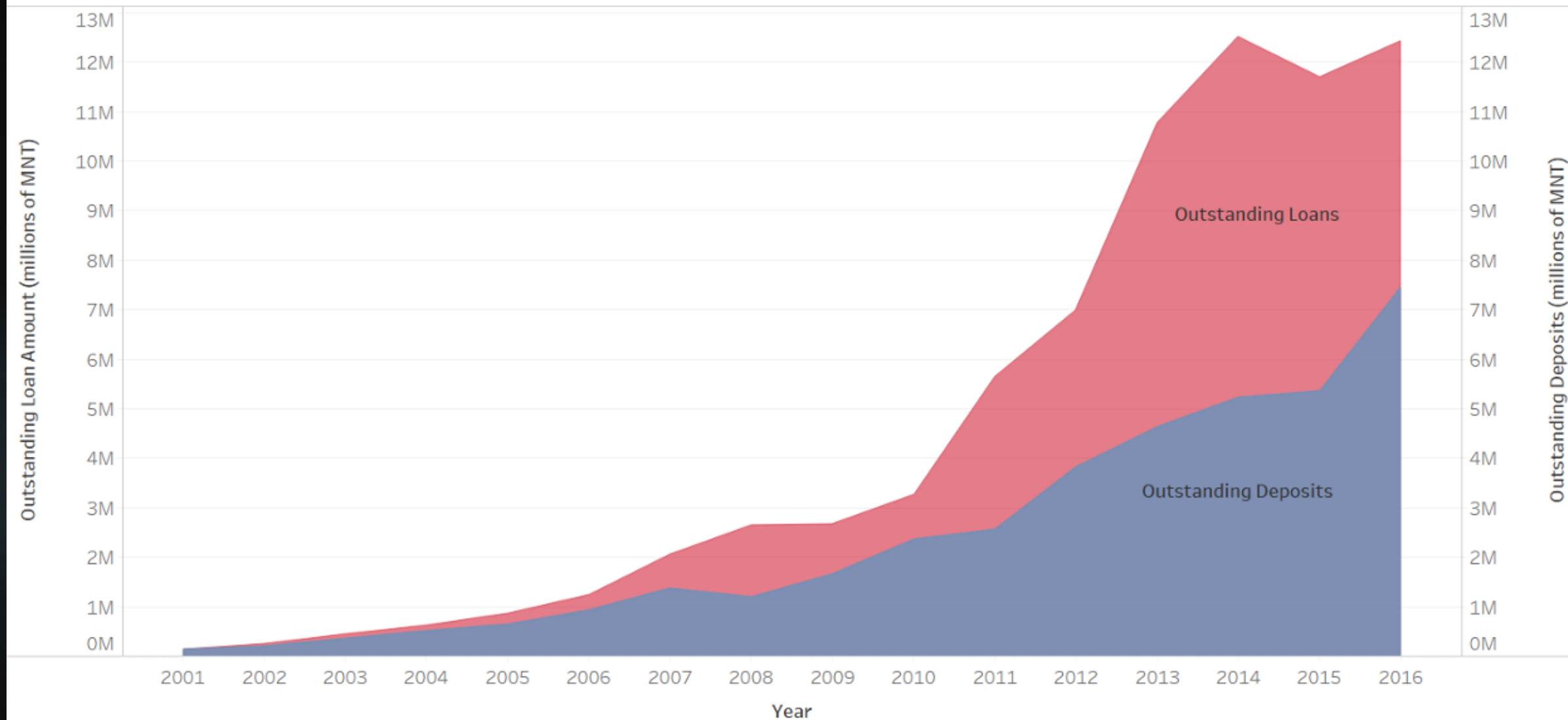
```
graph LR; A((Data was clean, just needed to reformat)) --> B((Interpretation is important))
```

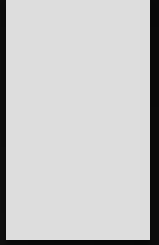
Data was  
clean, just  
needed to  
reformat

Interpretation  
is important

# Cleaning and Analysis

## Outstanding loan and deposit amount 2001-2016 (millions of MNT)

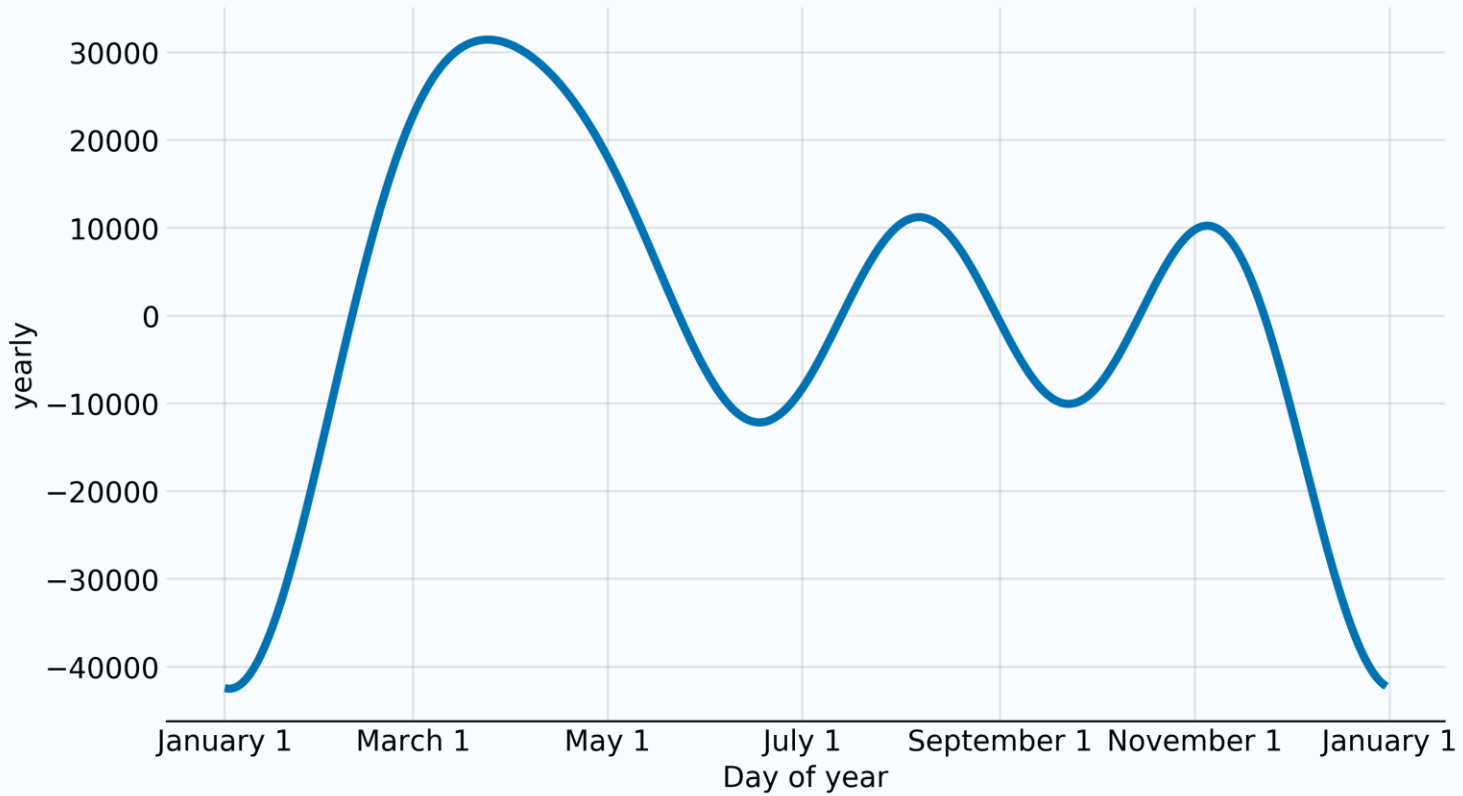




## Tools used:

- ▶ Excel – data cleaning and analysis
- ▶ Tableau

Pension Loan Cycle



Follow-up –  
Tsagaan Sar

A close-up photograph of a person's hands using a large knife to slice a piece of red meat on a white cutting board. The meat is being cut into thin, uniform strips. Several other pieces of sliced meat are scattered on the board, and a large pile of sliced meat is visible in the foreground, slightly out of focus. The background is a plain, light-colored surface.

# Meat Prices



2019 оны 1 сарын 4



**Б.БАДАМГАРАВ**

iKon.mn сэтгүүлч



Махны үнэ тэнгэрт хадаж, хомсдол үүсжээ!





MNB Монголын Мэдээ · Follow

## Мах экспортлогч ААН-үүд их хэмжээний мах нөөцөлсөн байгааг шалгалтаар илрүүлжээ...

Мах экспортлогч ААН-үүд их хэмжээний мах нөөцөлсөн байгааг шалгалтаар илрүүлжээ. #МонголынМэдээ #ЦагийнХүрд



Like



Comment



Share



619 · 292 Comments

# Data



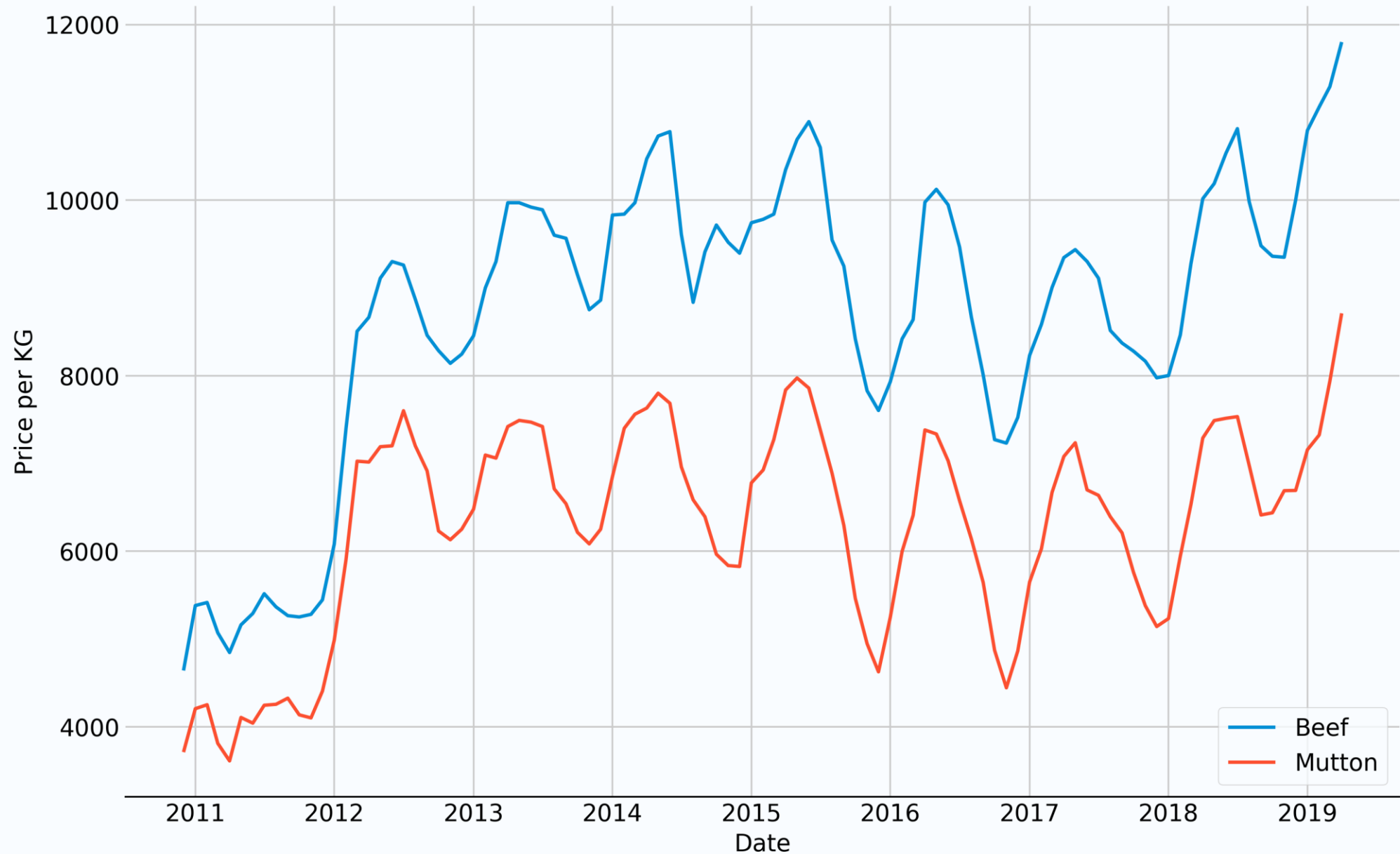
Luckily, the National Statistical Office tracks monthly meat price averages



Issue: so many kinds of meat!

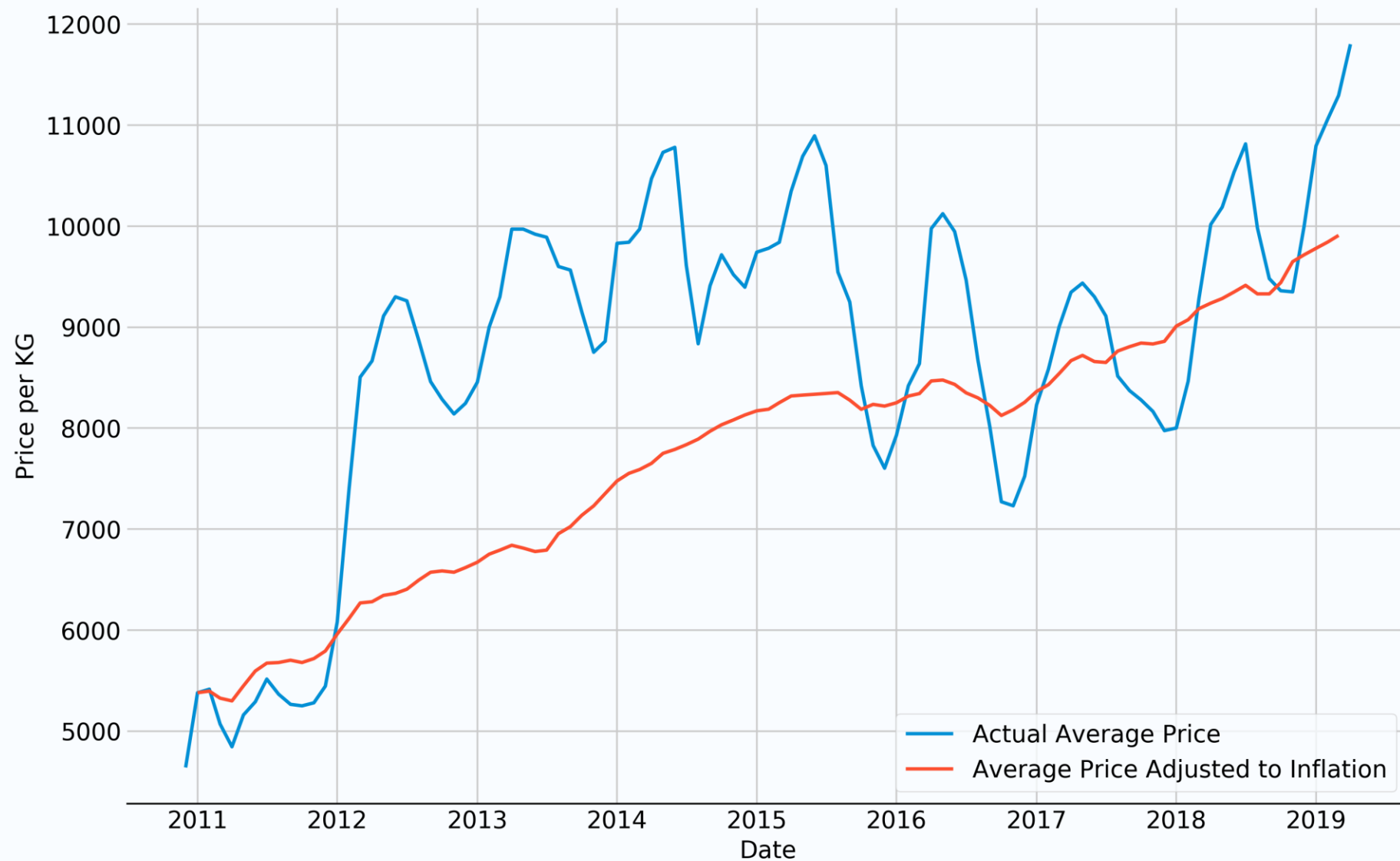
# Roller Coaster of Meat Prices

Dec 2011 to Apr 2019



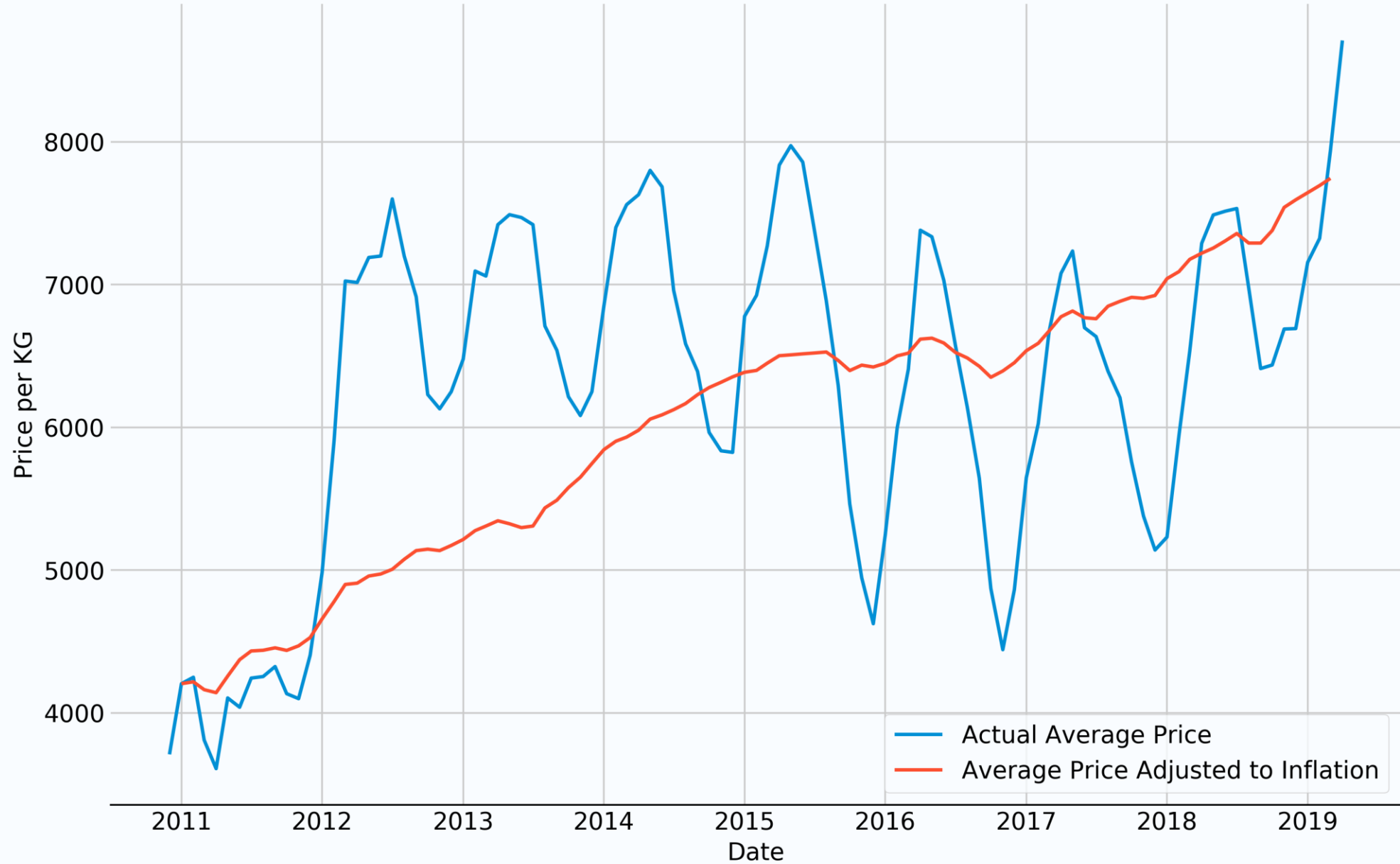
# Beef appears to be priced at a premium

Only 2017 saw beef prices tracking inflation



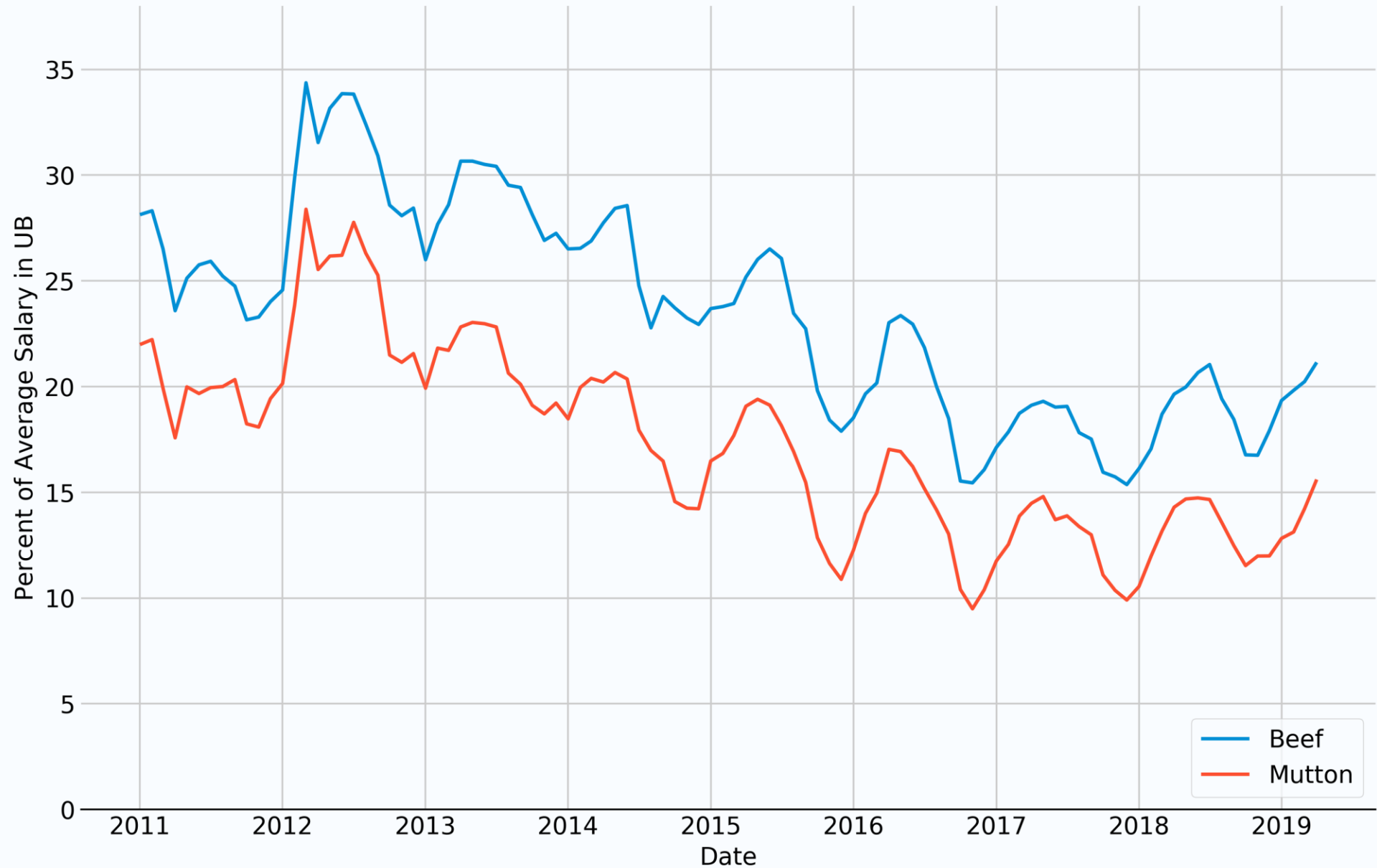
# Mutton prices rising above inflation trend

2016-2018 saw prices under inflation trend



# Meat as a share of income in going down

Based on average UB salary. Assumes 250g of consumption per day for three people.



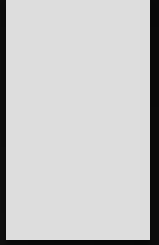


# Meat as a share of income in going down

Based on average UB salary. Assumes 250g of consumption per day for three people.





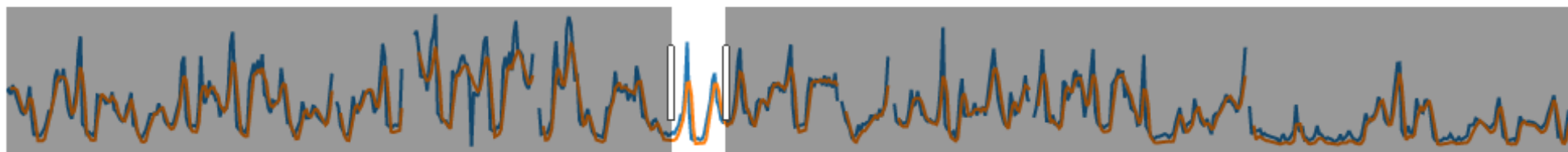
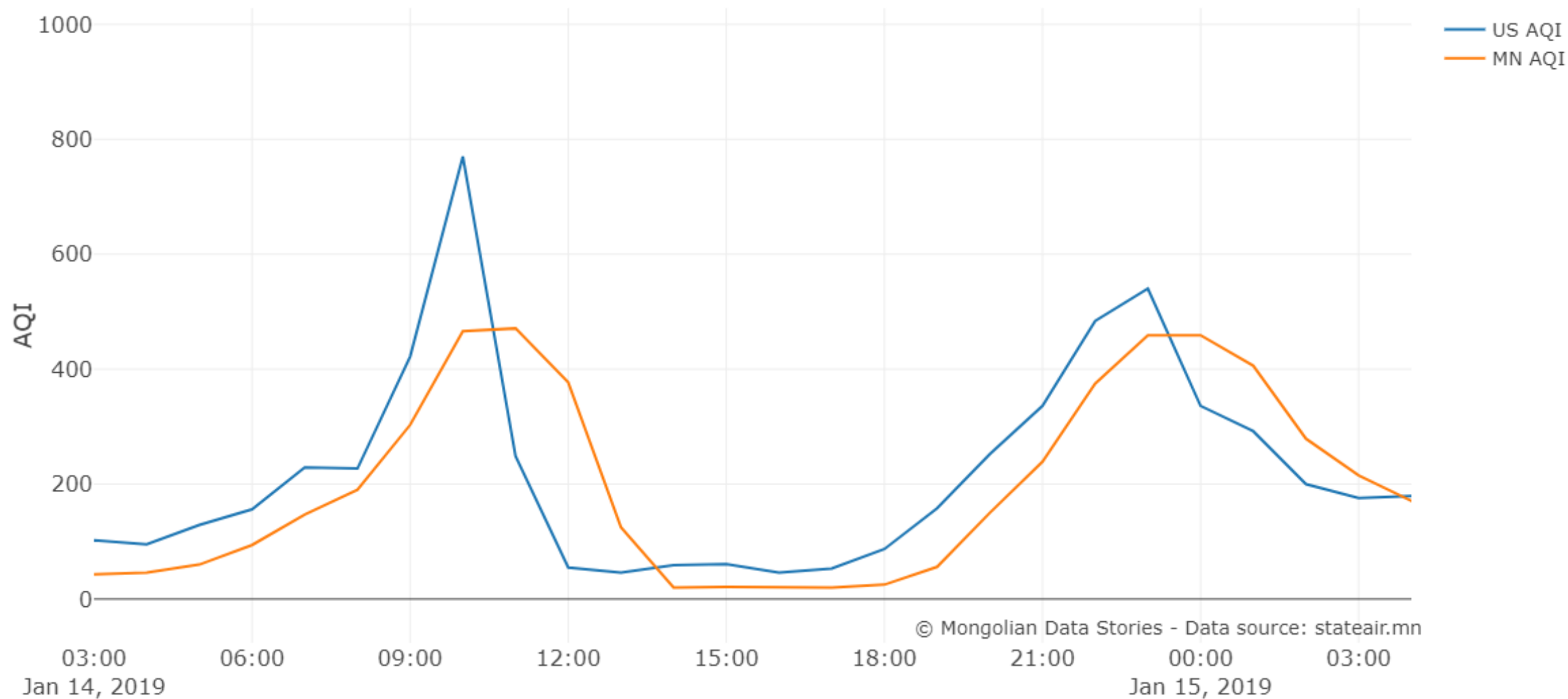


## Tools used:

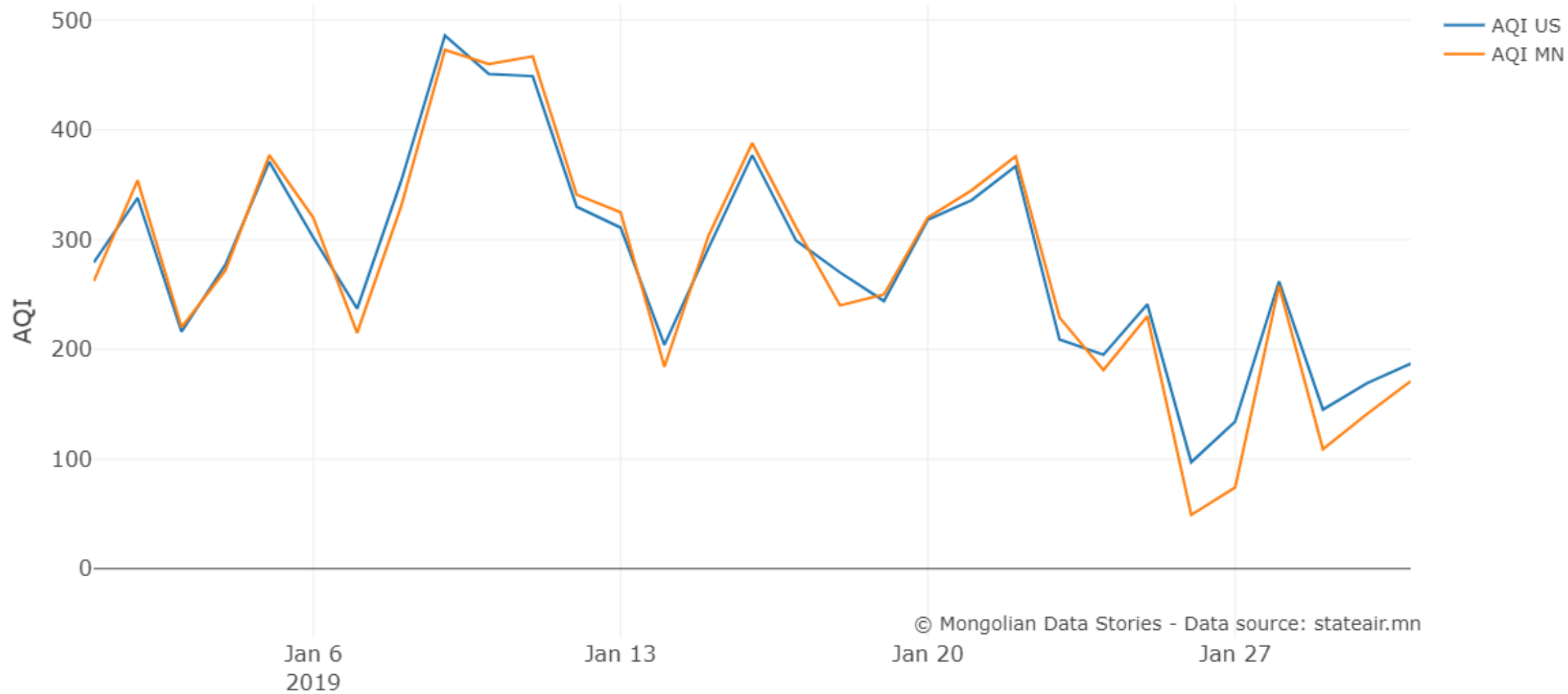
- ▶ Excel – data exploration
- ▶ Python – analysis and visualization

# Current Project – Air Pollution

## One Hour vs Three Hour Average - AQI



## AQI Daily Average



© Mongolian Data Stories - Data source: stateair.mn

# Data Analysis Process

