# Inspiration and skill sharing

For data journalism educators

Jonathan Stoneman and Sandra Foresti

Dataharvest 2024

# Presentation

- Who is here?

- What do you teach?

- What are the most rewarding part when teaching data journalism/investigative journalism?

    AND/OR

- What are the most difficult parts?

# Areas to talk about

- *What exercises do you have for the students/participants to better:*
  - *Understand data*
  - *Find the journalistic angle from data*
  - *Verification, source criticism and proofreading*
  - *Visualization and presentation of data journalism*

- *Other areas?*

# What should we focus on, in this session?

# Exercise to make the students to better understand structured data

# Young woman robbed

The 22 year old woman was walking on Övre Husargatan in the centre of Gothenburg on Tuesday morning, when she was robbed by a young man, dressed in black, threatening her with something that looked like a knife. He took her wallet, mobile phone and necklace.

This news item helps students understand structured data.

The students works in groups of two or three.

# Young woman robbed

The 22 year old woman was walking on Övre Husargatan in the centre of Gothenburg on Tuesday morning, when she was robbed by a young man, dressed in black, threatening her with something that looked like a knife. He took her wallet, mobile phone and necklace.

If the purpose is to have an overview of robberies in Gothenburg, about victims, offenders, time and place, what pieces of information is relevant to select?

# Young woman robbed

The 22 year old woman was walking on Övre Husargatan in the centre of Gothenburg on Tuesday morning, when she was robbed by a young man, dressed in black, thhreatening her with something that looked like a knife. He took her wallet, mobile phone and necklace.

**How do you transform the data in the text to machine-readable data?**

- Structure
- Rows and columns
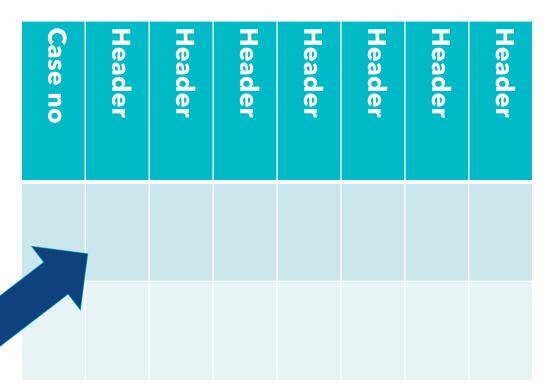- One variable at a time

# Young woman robbed

The 22 year old woman was walking on Övre Husargatan in the centre of Gothenburg on Tuesday morning, when she was robbed by a young man, dressed in black, threatening her with something that looked like a knife. He took her wallet, mobile phone and necklace.

| Case no | Header | Header | Header | Header | Header | Header | Header |
|---------|--------|--------|--------|--------|--------|--------|--------|
|         |        |        |        |        |        |        |        |
|         |        |        |        |        |        |        |        |

Take a spreadsheet and structure the pieces of information
Give each column a proper title

![University of Gothenburg logo]

Now, try to add THIS news item to your table.

*…you may need to change some headers for both items to fit…*

# Man robbed at Linnéplatsen

A stolen mobile phone and a black eye was the result when a man in the lower twenties was robbed at 4 am Tuesday morning. The robbery took place at Linnéplatsen in Gothenburg. The police is looking for two men, dressed in dark clothes

*Your spreadsheet could look something like this*

| Case | Day | Time | Place | Victim_sex | Victim_age | Things_robbed | Physical_injuries | Several_robbers? | Robber_sex | Robber_weapon |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | Tuesday | Morning | Övre Husar-gatan | F | 22 | Mobile phone, wallet, necklace | No data | No | M | Knife |
| 2 | Tuesday | 4 AM | Linné-platsen | M | Early twenties | Mobile phone | Black eye | Yes | M | No data |

# There is no perfect data
# Always problems:

What time in the morning?

How to count most frequently robbed object?

If they are a mixed gang?

| Case | Day | Time | Place | Victim_sex | Victim_age | Things_robbed | Physical_injuries | Several_robbers? | Robber_sex | Robber_weapon |
|------|-----|------|-------|-----------|-----------|---------------|-------------------|------------------|-----------|---------------|
| 1 | Tuesday | Morning | Övre Husar-gatan | F | 22 | Mobile phone, wallet, necklace | No data | No | M | Knife |
| 2 | Tuesday | 4 AM | Linné-platsen | M | Early twenties | Mobile phone | Black eye | Yes | M | No data |

What Tuesday?

The street is long…

How calculate average age?

How many?

Missing data

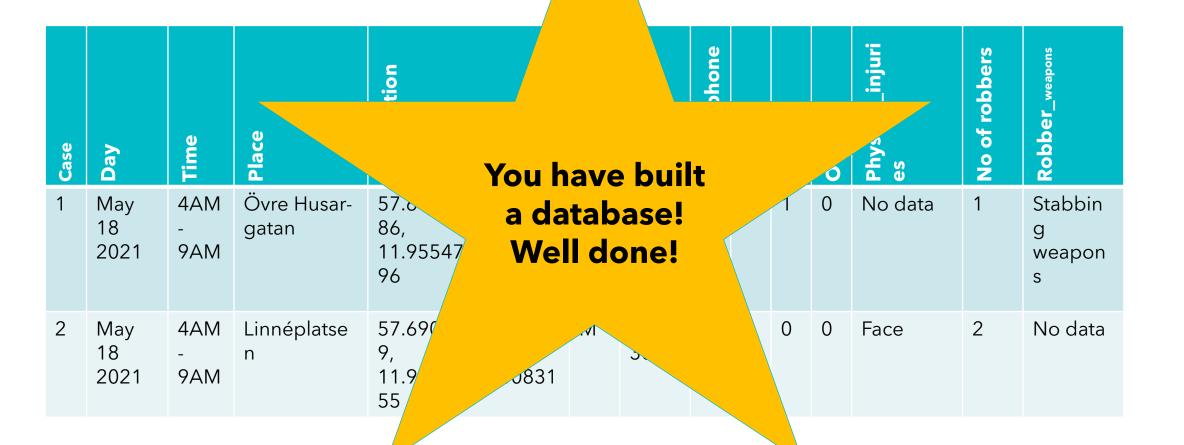After discussion:

Make a new version

# Make a new version

| Case | Day | Time | Place | Geolocation | Victim sex | Victim_age | Mobile phone | Wallet | Jewlery | Other | Physical_injuries | No of robbers | Robber_weapons |
|------|-----|------|-------|-------------|-----------|------------|--------------|--------|---------|-------|-------------------|---------------|----------------|
| 1 | May 18 2024 | 4AM-9AM | Övre Husar-gatan | 57.692356847784986, 11.955477787756996 | F | 20-30 | 1 | 1 | 1 | 0 | No data | 1 | Stabbing weapons |
| 2 | May 18 2024 | 4AM-9AM | Linnéplatsen | 57.69011474435299, 11.952312810083155 | M | 20-30 | 1 | 0 | 0 | 0 | Face | 2 | No data |

Find the date

Categorize when higher data types can't be used

Add data you need

One piece of information in each cell.

| Case | Day | Time | Place | ...tion | | ...hone | | | Phys_injuri es | No of robbers | Robber_weapons |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | May 18 2021 | 4AM - 9AM | Övre Husar- gatan | 57.6 86, 11.95547 96 | | | 1 | 0 | No data | 1 | Stabbin g weapon s |
| 2 | May 18 2021 | 4AM - 9AM | Linnéplatse n | 57.690 9, 11.9 55 | | | 0 | 0 | Face | 2 | No data |

**You have built a database! Well done!**

7. Imagine gathering all cases of robbery during one year in Gothenburg in your database.

*Write down three headlines possible to make, analyzing the data in your database.*

*Example: "[…] of the victims of robbery are female*

# Exercise to train to find journalistic angle in data (Combined files)

## Median income/inhabitants

| City area | Year | Median income before taxes/year |
|---|---|---|
| | | |

## Persons convicted for severe crimes

| Personal ID | Court act ID | Year | Crime | Sentence |
|---|---|---|---|---|
| | | | | |

## Pollution

| Place (lat/lon) | Date | Time | Substance | Level |
|---|---|---|---|---|
| | | | | |

## Causes of death

| Postal code | Year | Cause of death | Count |
|---|---|---|---|
| | | | |

## High school grades

| Name of school | Year | Subject | Average grade | Count of pupils |
|---|---|---|---|---|
| | | | | |

## Register of school

| Name of school | Address | Postal code |
|---|---|---|
| | | |

## Register of addresses

| Address | Postal code | City area |
|---|---|---|
| | | |

## Amount of traffic

| Place (lat/lon) | Date | Count of vehicles |
|---|---|---|
| | | |

## Opinion of own health

| City area | Year | Share who thinks they have good or very good health |
|---|---|---|
| | | |

## Unemployment 18-64 years

| City area | Year | Share of unemployed |
|---|---|---|
| | | |

## Election result

| City area | Year of election | Party | Share of votes |
|---|---|---|---|
| | | | |

## Weather at 12:00

| Place (lat/lon) | Date | Degrees Celsius | Rain since day before (mm) |
|---|---|---|---|
| | | | |

## Type of living

| City Area | Type (eg apartment, house) | Rental/Owned | Share of type in area |
|---|---|---|---|
| | | | |

## Food services

| Company ID-number | Address | Type (Café, restaurant..) |
|---|---|---|
| | | |

## Pre school

| Name of pre school | Address | Count if children | Count of children in queue |
|---|---|---|---|
| | | | |

## Real estate register

| Property_ID | Address | Owner_ID | Value |
|---|---|---|---|
| | | | |

## Hygiene controls in food service companies

| Company ID number | Date | Status (OK/Not OK) |
|---|---|---|
| | | |

## Reported crimes

| Place (lat/lon) | Date | Reported crime |
|---|---|---|
| | | |

# Exercise 1

- Groups of 2-3

- Find journalistic relevant hypothesis' by combining two or more of this hypothetic files

- Find at least three

- Try to identify a usable primary key. If there isn't any – how could you solve it?

# Exercise to train proofreading and Line by line

1

2  Staffanstorp municipality

3  The reporter is Sophia Djiobaridis.

4  I'm on my way from Staffanstorp in Skåne - to Cannes on the French Riviera - m
5  same journey that some politicians and officials made in March this year. The r
6  follow the invoices and receipts submitted by the heads of the municipalities t
7  done down there and where the tax money is going.

8  The very first receipt from day one is an early taxi ride from Staffanstorp to Co
9  cost was SEK 1907. There are actually five representatives from the municipalit
10  does not join until a few days later.

11  Reporter: And now we're looking for the next time you swipe the card and it's
12  Sommelier Bar & Bistro and it's 7.38am. I would like to have one of each that is
13  receipt.

14  Waiter: Yes.

15  Reporter: One of each.

16  Waiter: Yes. A breakfast buffet?

17  Reporter: Yes.

18  Waiter: Also a Carlsberg beer?

19  Reporter: Yes.

20  Waiter: How about an Akvavit?

21  Reporter: Exactly

## LE SOMMELIER
### BAR & BISTRO
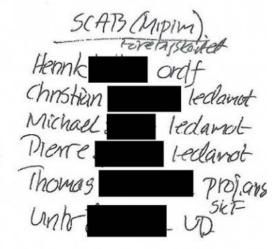
Copenhagen Airport
Phone +45 3248 3142

Invoice for
Table number 2

```
6 *  149.00 DKK/Stk
Morgenbuffet                      894.00 M
  5 *   59.00 DKK/Stk
Carls. fad mellem                 295.00 M
Juice appelsin 50cl                50.00 M
  5 *   45.00 DKK/Gl 3cl
Gammel dansk 3cl                  225.00 M

Subtotal EUR                      199.46
Subtotal DKK                     1464.00
Fadøl 50cl+3cl bitte              -45.00

Subtotal EUR                      193.32
Subtotal DKK                     1419.00
Creditcard DKK                   1419.00

net turnover                     1135.20
25.00 %-vat M                     283.80
Grossturnover                    1419.00
```

Ticket  Date  Time  Oper  Term
326702 090313 0938 04203 0425
Vat no.: 71750515

Invoice No 3184504

SCATB (Mipim)
Foretajskodet

Henrik ███████ ordf
Christian ███████ ledamot
Michael ███████ ledamot
Pierre ███████ ledamot
Thomas ███████ projans
sk F
Unh ███████ UD

Mipim

Eyexon  42
KBH LUFTHAVN
2770 Kastrup
Tlf. +45 32483142
CVR nr. 71750515

2013-03-09                09:21

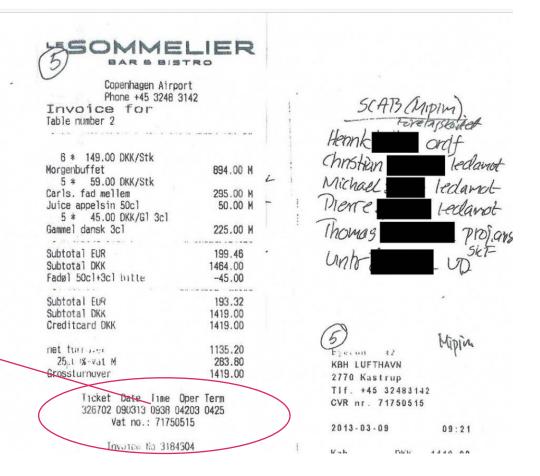Køb          DKK      1419.00

# Manuscript Mission investigate

1

2    Staffanstorp municipality

3    The reporter is Sophia Djiobaridis.

4    I'm on my way from Staffanstorp in Skåne - to Cannes on the French Riviera - my goal is to make the
5    same journey that some politicians and officials made in March this year. The method is simple - I
6    follow the invoices and receipts submitted by the heads of the municipalities to see what is being
7    done down there and where the tax money is going.

8    The very first receipt from day one is an early taxi ride from Staffanstorp to Copenhagen airport. The
9    cost was SEK 1907. There are actually five representatives from the municipality, but one of them
10   does not join until a few days later.

11   Reporter: And now we're looking for the next time you swipe the card and it's something called
12   Sommelier Bar & Bistro and it's 7.38am. I would like to have one of each that is ordered on this
13   receipt.

14   Waiter: Yes.

15   Reporter: One of each.

16   Waiter: Yes. A breakfast buffet?

17   Reporter: Yes.

18   Waiter: Also a Carlsberg beer?

19   Reporter: Yes.

20   Waiter: How about an Akvavit?

# Exercise to find the story they want to visualize