EY Data & Analytics Case
Data Analysis puzzle solving

2021-11-11

The better the question. The better the answer.
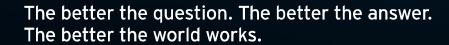The better the world works.

EY
Building a better
working world

A Formel 1 driver has been found dead in Nagoya, Japan...

The F1 driver was killed in Nagoya: **Sebastian Vettel** found dead after a pre-race dancing night.

His rental car was found burning next to the Port Bluebonnet Wildflower Garden in Nagoya.

The burning is estimated to have occurred in the night of **October 12th 2019**.

The body of the victim was found in the trunk of the burning car.

… and the Japanese police suspects that another Formel 1-driver might be guilty of the murder…

- At the night of the murder, drivers from all over the world were in Nagoya to have fun before the race

- The competition between the teams is intense and the drivers always look for strong emotions

- The excellent detectives of the Japanese police suspect, given the above, that one of the drivers may be the murderer

EY

... using one or more of these weapons, which has all ben found at the crime scene

The police need the help from your team to analyze the data to figure out and find proof on who commited the crime

Task: Find out the answers to

- ▸ Weapon used to commit the murder?

- ▸ Where the murder took place?

- ▸ When he was murdered?

- ▸ Who murdered him?

- ▸ Why did the defendant commit the crime?

Present

3 min presentation when you present your case, including a timeline around the time of the murder as well as answers to above questions. Please also include a short description of how you solved the case.

EY

tweets.xlsx and nrx.xlsx

tweets.xlsx

▸ Provides information about tweets made by each driver

▸ Contains 4 columns
- ▸ ID: identifier of the tweet
- ▸ Created at: datetime of created tweet
- ▸ Text: tweet content
- ▸ TwitterName: user

nrc.xlsx

▸ Provides information about the sentiment of words and can be used for sentiment analysis

weapons.csv and murderdata.csv

weapons.csv

▸ Contains information about chemical samples of various murder weapons in previous crimes seen by the Amsterdam Police Department

▸ The information includes 5 types of weapons:
  ▸ Revolver
  ▸ Knife
  ▸ Lead pipe
  ▸ Wrench
  ▸ Candlestick

murderdata.csv

▸ Information about two samples found on the victim, containing different levels of the following chemical components:
  ▸ RI, Al, Mn, Pb, Si, K, Cu, Ag, Fe

EY

GPSdata.csv, 112calls.csv and PhoneData.csv

GPSdata.csv

- The burnt car was leased from a Japanese company that required telematics to be installed. This telematics unit amongst others was gathering the following data in the days before the murder:
  - Time
  - Latitude
  - Longitude

PhoneData.csv

- Information about pgone history of the leading suspects (drivers) was found, which contains:
  - ID (name)
  - Type
  - Longitude
  - Latitude
  - Time (of call or text)

112calls.csv

- The police department supplied with the logs of emergency calls from the days around the murder, containing:
  - Latitudes (N-degree)
  - Longitudes (E-degree)
  - Emergency type
  - Date
  - Time

EY

fraud.csv, fraud_target.csv and suspects.xlsx

fraud.csv

▸ Contains the transaction records of the suspects in the past that were flagged as possible fraud which can be used for training purposes.

fraud_target.csv

▸ Contains the transaction records of the suspects of the day before the murder. This can be used to find the prime suspect(s).

suspects.csv

▸ Contains information about all suspect drivers:

  ▸ Name
  ▸ Twitter account
  ▸ Bank account
  ▸ Formel 1 team

Useful tips along the way

- Keep in mind that different steps in your investigation lead to new insights
    - E.g. where and when are evidently connected

- Feel free to use the tool of you choice. Examples of tools are Python, R, Excel/VBA, Tableau, Spotfire and Power BI

- Your team can use the *words* dataset if you want to do a sentimental analysis

- You can use e.g. the Scikit-learn library in Python if you want to build a Machine Learning model useful for data analysis

EY

The drivers are leaving Nagoya soon, so make sure to be finished with the investigation by then...

Solve the case!

You will have **2 hours** to solve the murder mystery, excluding a 10 minute break half way.
Use your time wisely!

Present your solution

You will get **5 minutes** to present your solution. Keep in mind that the presentation should be in English.

Keep it clear

Keep in mind that the case should be presented in a clear and logical way, including what your solutions are and how you derived them.

Send it in

Make sure to send your presentation to mathias.johnson@se.ey.com before the flight takes off!

GOOD LUCK!