

DATA COLLECTION & EXPLORATION



János Divényi

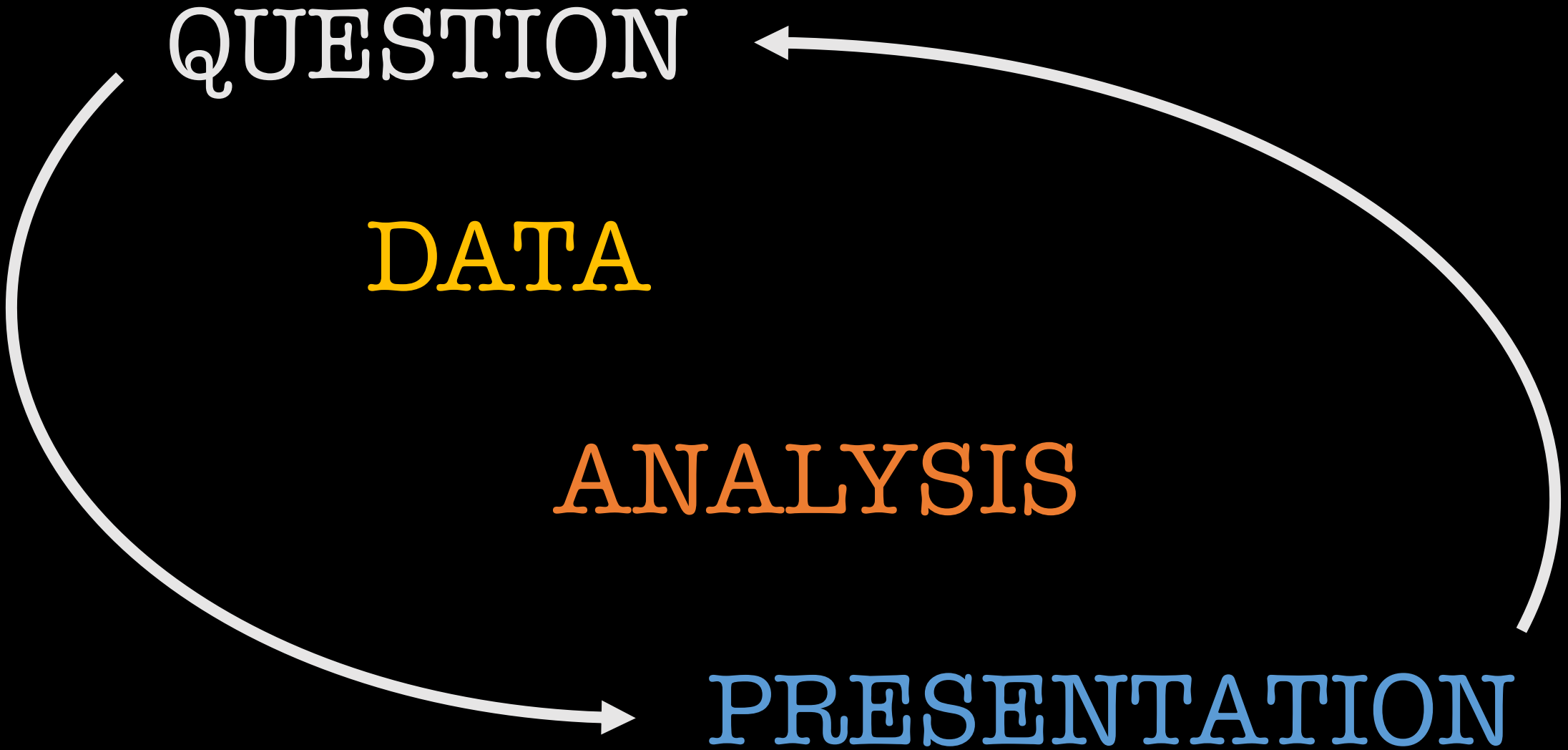
ELTECON DATA SCIENCE

QUESTION

DATA

ANALYSIS

PRESENTATION



DATA

DATA COLLECTION

eurostat

KSH

MNB

World Bank

Where can I find large datasets open to the public?



Answer



Follow · 5.9k



Request



22



<https://www.quora.com/Where-can-I-find-large-datasets-open-to-the-public>

experiment

survey

experiment

survey

scrapping (rvest)

WHERE DOES YOUR
DATA
COME FROM?

SAMPLING

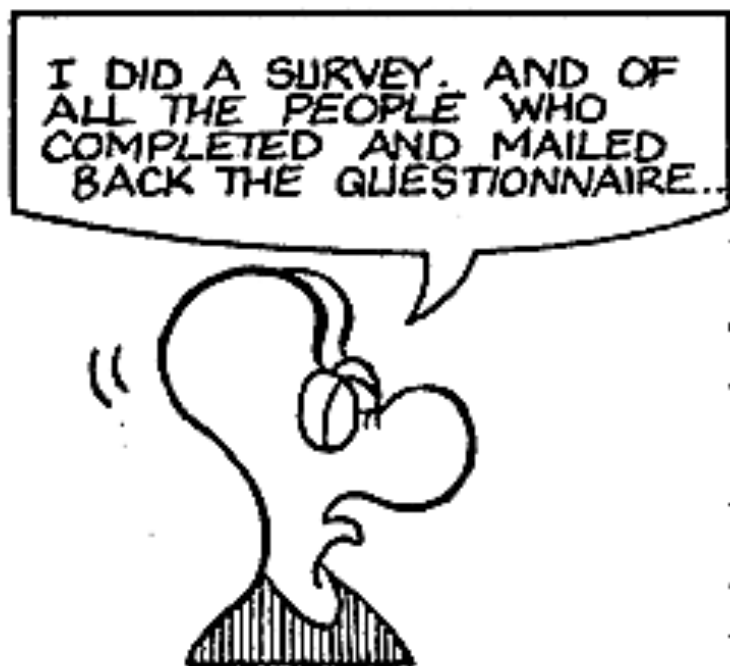
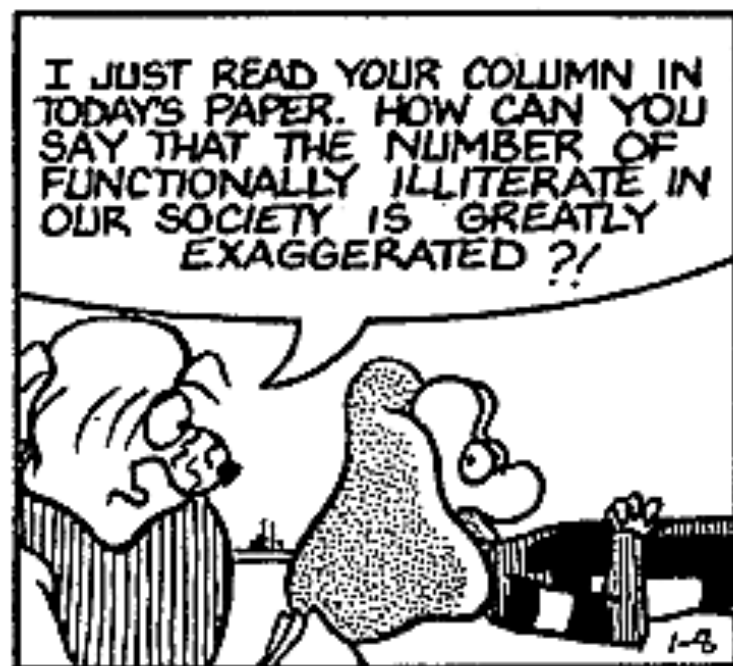


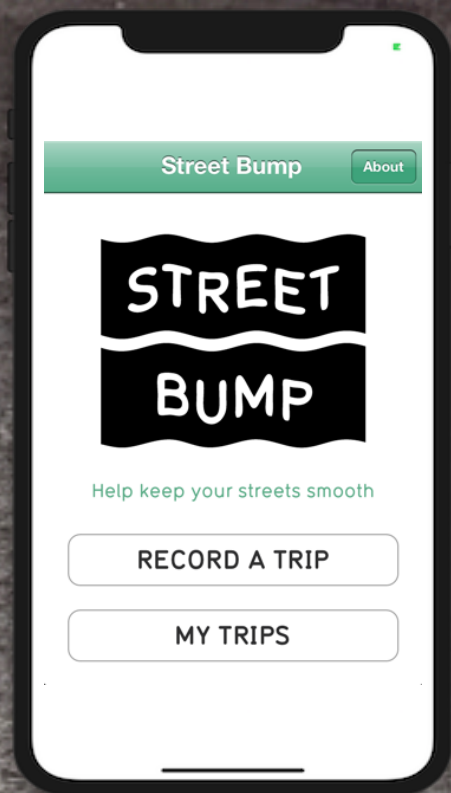




<https://boredbug.com/wp-content/uploads/2015/08/onesecondbeforedisaster.jpg>

THE MICE SQUAD





Street Bump

About

STREET
BUMP

Help keep your streets smooth

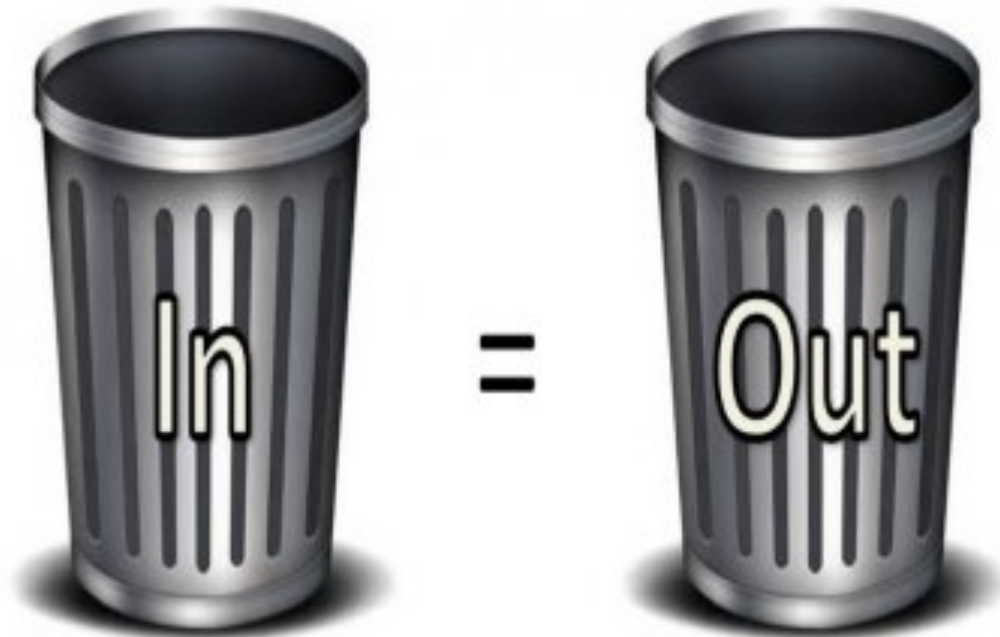
RECORD A TRIP

MY TRIPS

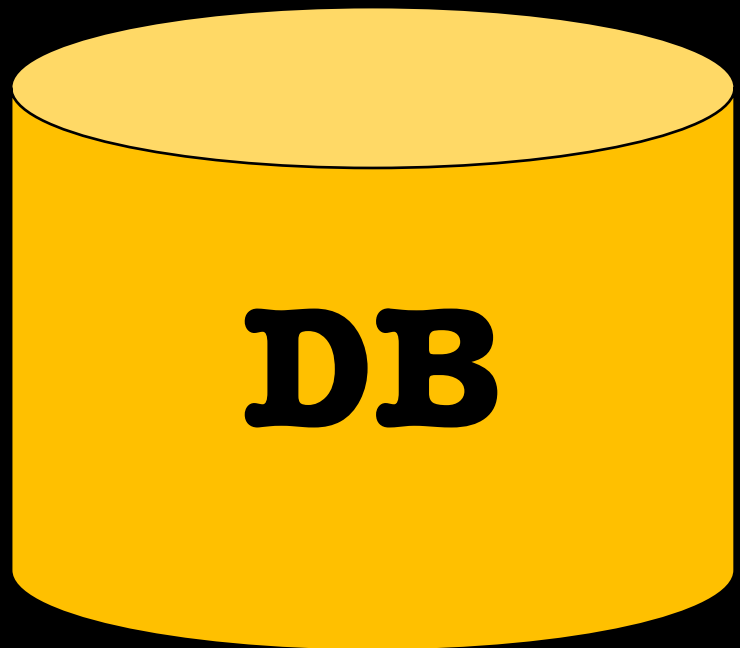
IMPLEMENTATION

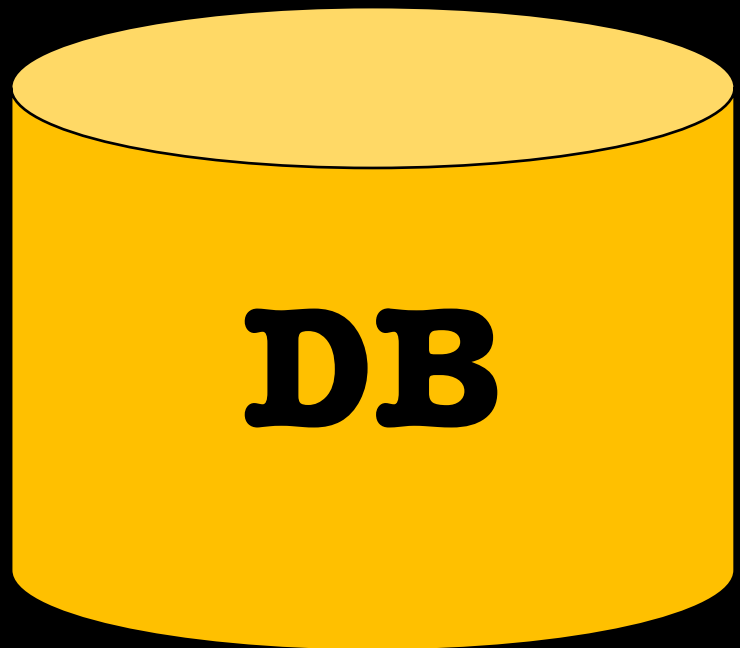
GIGO

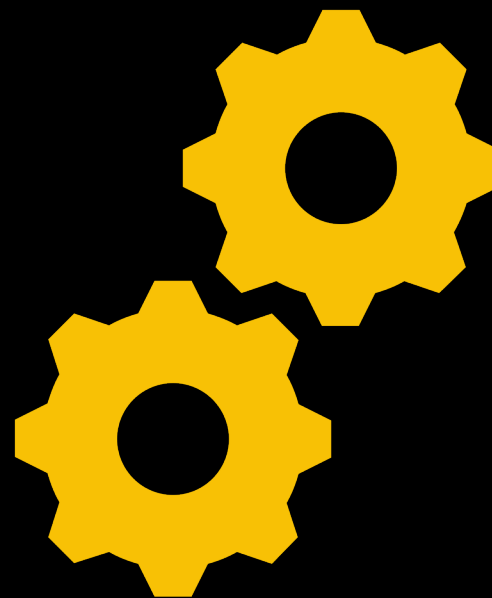
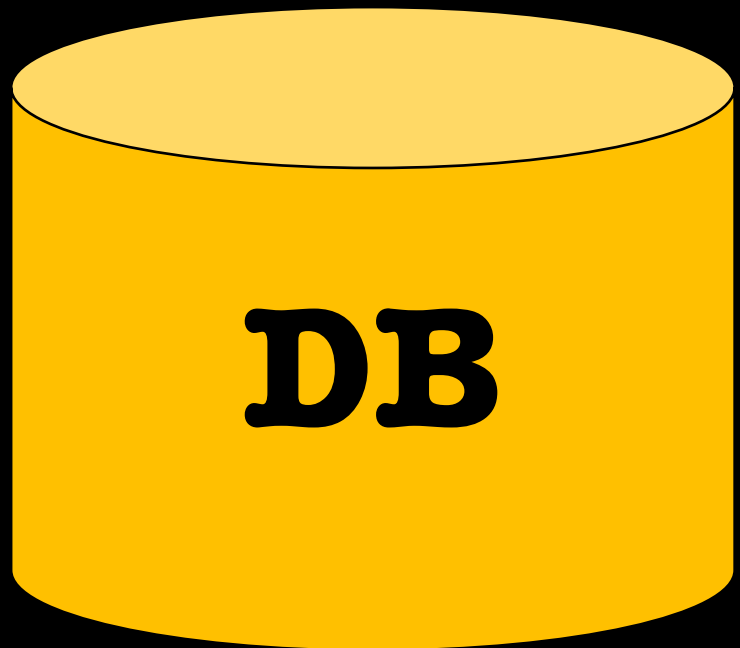
GIGO



DATA ACCESS







DT[**i**, **j**, **by**]

WHERE **SELECT** **GROUP BY**

Take table **DT**,
subset rows using **i**,
then calculate **j**,
grouped by **by**



www.w3schools.com/sql/exercise.asp

EXPLORATORY

DATA

ANALYSIS

KNOW YOUR
DATA

outliers

distributions

missing values

duplicates

errors

MAKE

CONSCIOUS DECISIONS