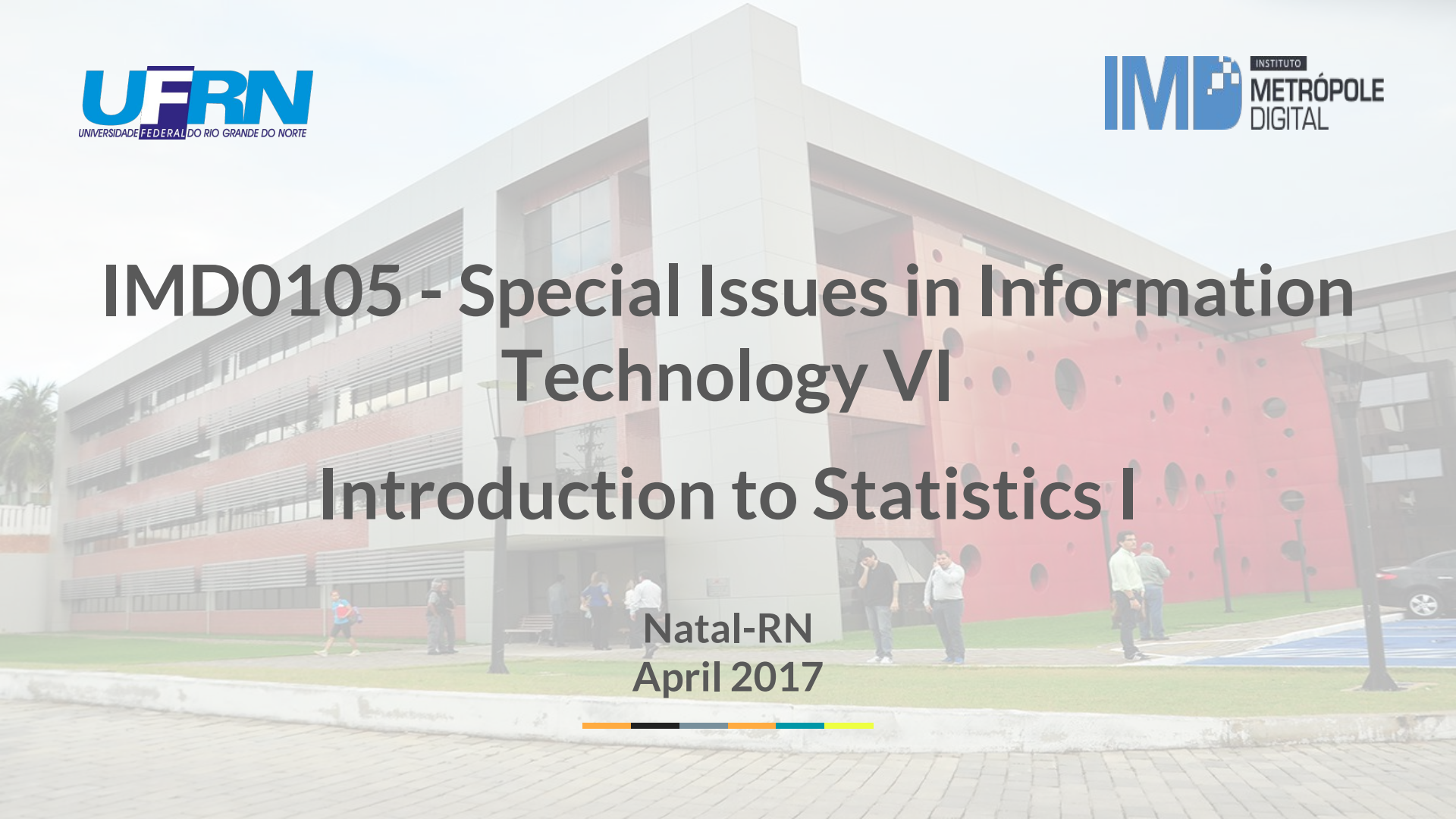


IMD0105 - Special Issues in Information Technology VI

Introduction to Statistics I

Natal-RN
April 2017

Agenda

- Central tendency measurements
  - Frequency
  - Histogram
- Practical
  - Open data Natal

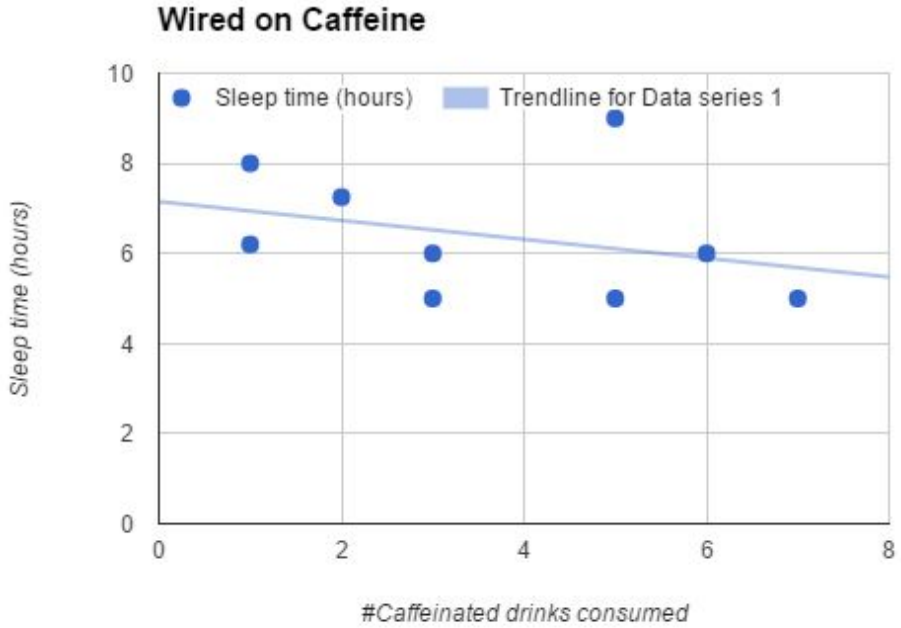
Previously on last class (...)

Caffeine

Quiz

Room: 4PSX6ZWHG

socrative
by MasteryConnect



Quiz
Room: 4PSX6ZWHG

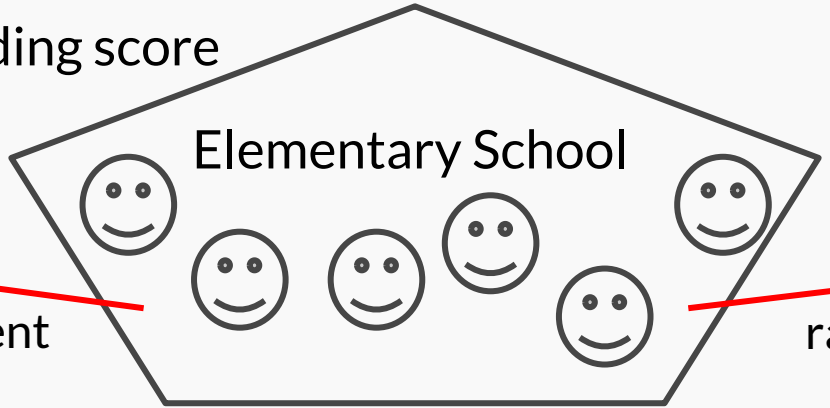
English Teaching

Two ways of teaching them English

1

78% reading score

English only



random assignment

22°C

Students (portuguese-speaking, english-learners)

55%

2

Bilingual

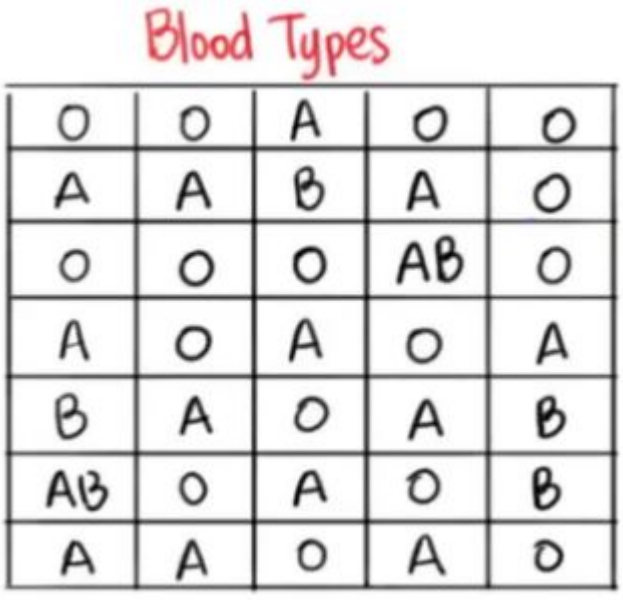
random assignment

35°C

Introduction to Statistics

Blood Types

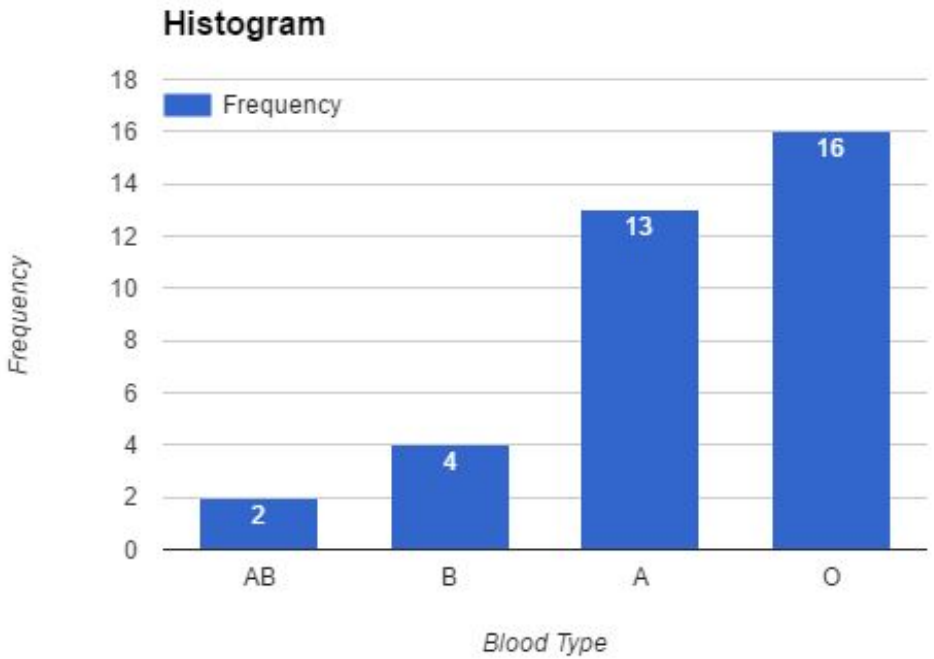
Studying about frequency



Blood Type	Freq.	Perc.	%
O			
A			
B			
AB			
Σ			

Blood Types

Studying about frequency



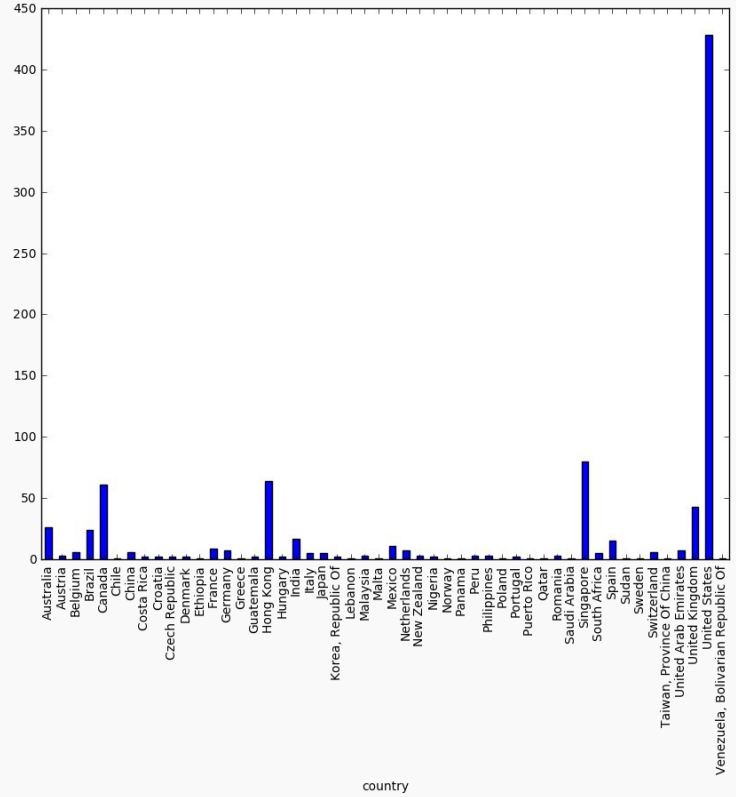
Where students are from?

Studying about frequency

	id	industry	country	city
0	10001	Agriculture	United States	Davis
1	10002	Arts & Education	Australia	Perth
2	10003	Arts & Education	Austria	Lieboch
3	10004	Arts & Education	Brazil	São Paulo
4	10005	Arts & Education	Canada	Georgetown
5	10006	Arts & Education	Canada	Hamilton
6	10007	Arts & Education	Canada	Milton
7	10008	Arts & Education	Canada	Mississauga
8	10009	Arts & Education	Canada	Regina
9	10010	Arts & Education	Canada	Toronto

Where students are from?

Studying about frequency



Download:

- WhereAreStudentsFrom.ipynb: https://goo.gl/m5NMyP
- Students_raw.csv: https://goo.gl/broV3K

Geocoder, folium, open data

DADOS ABERTOS
Prefeitura do Natal

http://ckan.imd.ufrn.br/

- Histogram/bar figure reporting the number of "municipal schools" and "health units".
- Geocoder figure about those units.
- A heatmap figure considering the number of employees in each unit.

Tips

```
# Import pandas
import pandas as pd

# Assign url of file: url
url = 'http://ckan.imd.ufrn.br/dataset/bfe9529d-5a43-49e1-b855-8f1caff9f123/re

# Read file into a DataFrame and print its head
df = pd.read_csv(url, sep=';')
```