

An aerial photograph of a winding asphalt road that curves through a dense, green forest. The road is light-colored and contrasts with the dark green trees. The perspective is from above, looking down at the road as it snakes through the landscape. The background is a solid orange color that serves as a backdrop for the text.

# Practical Assessment

Business Analytics - Text Mining

Tiago Otto Rodrigues

**P.PORTO**

PORTO  
EXECUTIVE  
ACADEMY

**P.PORTO**  
ISCAP

# The Goal

Welcome to our practical assessment to Text Mining!

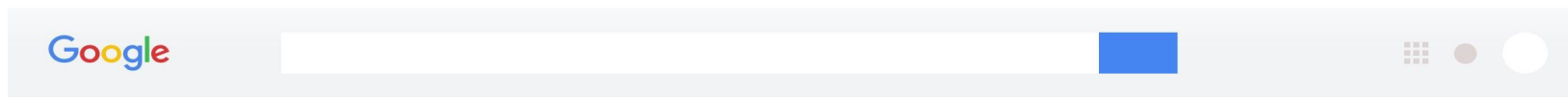
This challenge can be done in groups and represents 50% of the final grade.

The challenge is presented as a role play practical exercise, divided in different tasks.

You are about to become Rachel, a Data Scientist joining the financial department of US Airways.

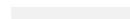
Have fun!

# Context

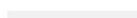
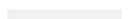
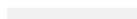
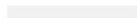


**From:** Justin (HR Director)

February 20th, 2023



**To:** Rachel



Good morning Rachel!

Welcome to the US Airways family.

Everyone is looking forward to finally meeting you, especially Michelle (our Chief Digital Officer). The early incorporation of a Data Scientist into the team has been one of her priorities. She considers you a key element to transform US Airways into a “data driven” company. She has a lot of ideas for your initial months with us. I’m sure you will be hearing from her very soon.

I wish you a great start in the new job. If there is anything I can help with, please let me know.

See you soon,

Justin





**Analysis**

# Task 01

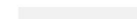
# Context: Analysis

Google

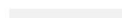
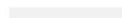
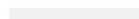


**From:** Michelle (Chief Digital Officer)

February 21th, 2023



**To:** Rachel



Hello Rachel,

First of all, welcome to US Airways!

I have already requested access for you to the company's data warehouse, our data collection department has been monitoring all tweets about our company and competitors ([link](#)).

I suggest you start looking at this dataset to familiarise yourself with the business.

See you soon!

M

# Context: Analysis

Google

**From:** Michelle (Chief Digital Officer)

February 21th, 2023

**To:** Rachel

Hello Rachel,

Now that you are more familiar with our data. I have a few questions I would like you to help me with:

- How is our company's twitter popularity in relation to our direct competitors?
- Can you identify what issues we should focus on to try to improve customer satisfaction?

Thanks in advance!

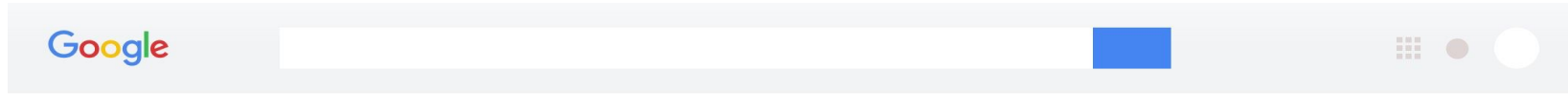
M



# Classification

# Task 02

# Context: Classification

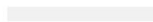
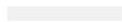
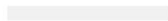


**From:** Michelle (Chief Digital Officer)

February 21th, 2023



**To:** Rachel



Hello Rachel,

Your previous analysis was super helpful!

Unfortunately, due to high costs, it is impossible for us to continue to label every single tweet about aviation. To solve this problem, do you think it is possible to build a Data Science model to automatically classify the sentiment of future tweets? Can you predict how confident we can be in this model to use it instead of manual annotation?

Thanks in advance!

M