Capgemini

DATA SCIENCE CONSULTING Session 1



Capgemini Invent, a leader in digital & data transformation



Strategic focus,

of our firm since 2012, now standing for 50% of our project portfolio



Leader on the market,

with award winning thought leadership, partnership with MIT, and world class recognition



World class footprint,

16 offices, 3,000 consultants, coverage of more than 80% companies of CAC 40 and DAX 30 at CxO level

Acknowledgements of business expertize



Gartner

Capital



Full spectrum digital specialist, Best-of breed provider - 2016

Leader position in the Magic Quadrant for Business Analytics Services 2017

N°1 consultancy for digital, data and transformation - 2016

Top 3 worldwide thought leadership consultancy - 2016



Objectives of the case study





Handle a business problematic associated to data

Increase both knowledge and skills on these topics



Learn how to determine & realize the required analysis

Handle a data project from the beginning to the end



Understand the strategic & transformation stakes

Qualify and quantify the associated stakes



Grasp the consulting aspects

Learn how to manage these kinds of projects



Presentation of the client





Multinational hospitality group founded in 1967 in France



Covers variety of segments: luxury, midscale and economy Subsidiaries in events organization and digital hospitality (catering, coworking...)



Operates in 100 countries – 4800 hotels and 280 000 employees worldwide



Rebranding strategy since 2011 on existing assets and new acquisitions



« Life styled for you » - Complementarity between group brands for better customer experience







Focus of case study: Novotel Canary Wharf



Rebranding of Bokan 39

The bar & restaurant above our Novotel in Canary Wharf, London with the goal of increasing its number of customers

- Use data to make the right choices for the rebranding of the space
- Suggest a new branding and type of cuisine
- Estimate the cost, benefits and breakeven point



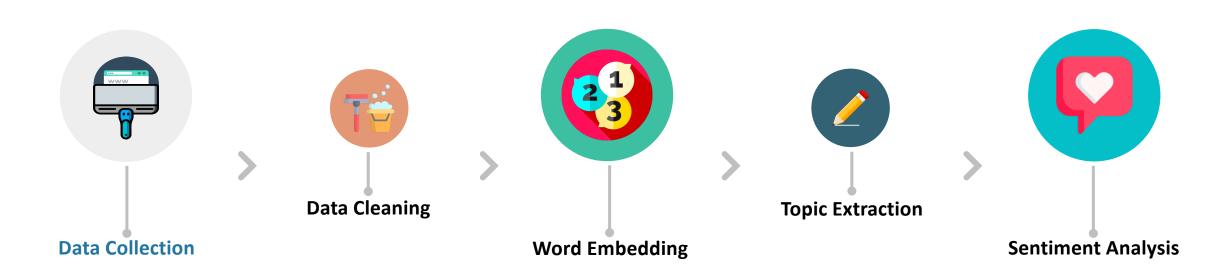






Data pipeline







Data Channels



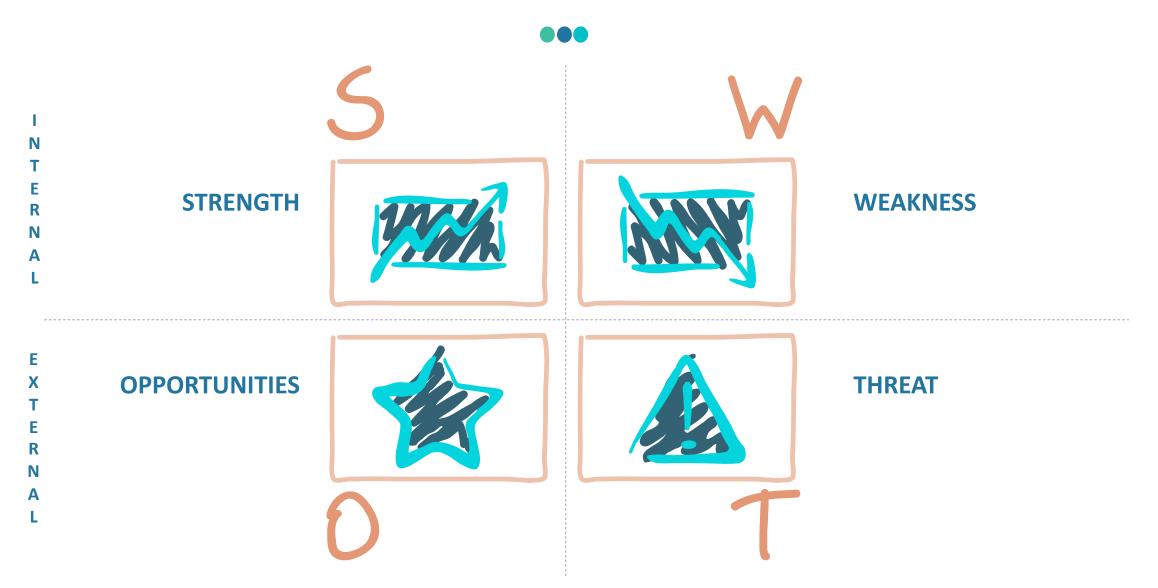






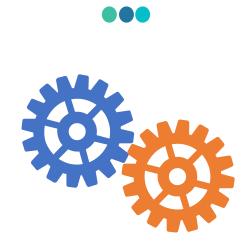


SWOT Analysis, what's it about?



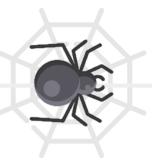


Two main blocks of Scraping



Parsing

Breaking down the scraped data into smaller bits to understand it



Crawling

Going through specified website and related links, to get raw data (source code)



Scraping's Main tools



BeautifulSoup

HTML parser in Python, which transforms HTML code into a tree object usable with Python to extract content like text, images' URLs...

h Scrapy

Python package combining a crawler AND a parser.

Selenium

Tool allowing to automate web browsers, mainly used in scraping to get data which is written in JavaScript through AJAX requests.

RVest

HTML parser like
BeautifulSoup,
but for R.
Great
documentation
online.



In this course we will focus on **Scrapy** exclusively.