eleven – data augmented proposal challenge

To the attention of the Data Science
for Business' master students

January 31st, 2022

HEC PARIS

ÉCOLE POLYTECHNIQUE
UNIVERSITÉ PARIS-SACLAY

eleven
strategy consultants
London • Paris • Madrid

AGENDA

HEC PARIS

eleven
strategy consultants

eleven is Europe's first ever specialist strategy firm specifically founded to accompany clients' transformation through the AI and digital revolution, thanks to a unique combination of strategy perspective and hands-on approach

DIGITAL & AI STRATEGY SPECIALIST

revolution is unlocking new, untapped value-creation opportunities

STRATEGY X HANDS-ON POSITIONING

supporting c-level executives and organizations from strategic ignition to project scale-up and

A UNIQUE BLEND OF SKILLS

and data science ones, thanks to its 50+ consultants that master the continuously evolving

DISTINCTIVE ENTREPRENEURIAL MINDSET

us to provide issue-based methodologies centered on your business needs
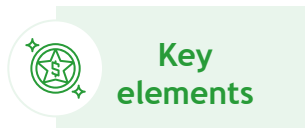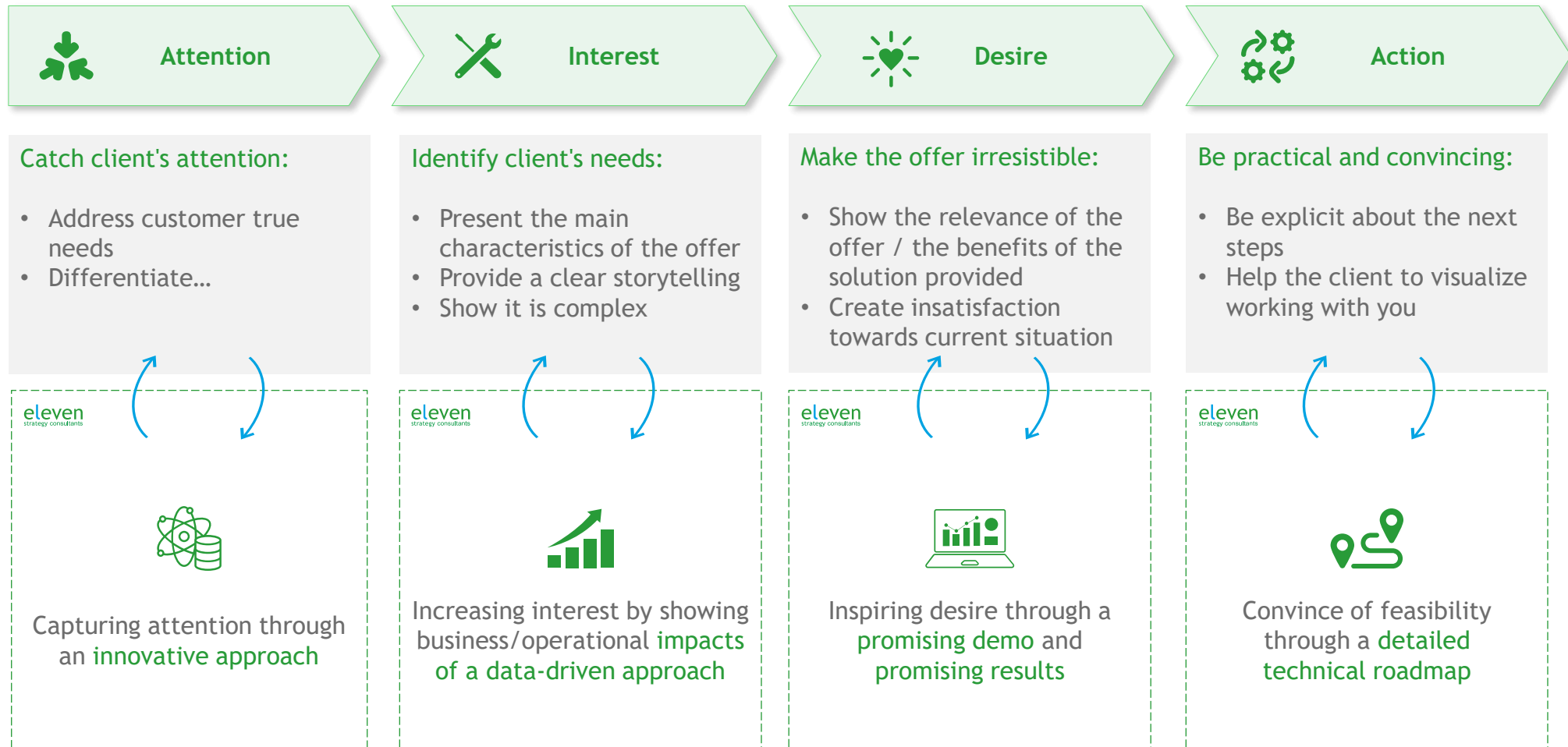
CSR AT THE CORE OF OUR DNA

Our approach enables our clients' projects to meet CSR expectations. Both financial and CSR impacts are tracked, proven and reported thanks to our mastering of digital and AI levers

AGENDA

Based on a major sales concept: AIDA, eleven relies on its unique expertise to provide innovative solutions to clients' problems

Attention	Interest	Desire	Action

Catch client's attention:

- Address customer true needs
- Differentiate...

Identify client's needs:

- Present the main characteristics of the offer
- Provide a clear storytelling
- Show it is complex

Make the offer irresistible:

- Show the relevance of the offer / the benefits of the solution provided
- Create insatisfaction towards current situation

Be practical and convincing:

- Be explicit about the next steps
- Help the client to visualize working with you

---

eleven strategy consultants

Capturing attention through an **innovative approach**

eleven strategy consultants

Increasing interest by showing business/operational **impacts of a data-driven approach**

eleven strategy consultants

Inspiring desire through a **promising demo** and **promising results**

eleven strategy consultants

Convince of feasibility through a **detailed technical roadmap**

Key elements

✓ **Data enriches the discourse** and makes it more effective at each stage of a commercial proposal

Your mission: structure a data-augmented commercial exchange

You are expected to deliver the following:

- Your code used to solve the use case, including any relevant dashboard, data analysis, model training, etc. This code should be readable by anyone, with a clear *Readme*, *requirements*, potential *quickstart.py*, etc.

- A "client-ready" PowerPoint presentation of your work, structured as follows:

1

Context: what do you understand from the project? Why is the context very complex?

- ✓ Present your understanding of the situation
- ✓ Include the process, the needs, the potential ROI, etc.
- ➢ c. 2/3 slides

2

Key Success Factors: what would be the reason of your success?

KSF	WHY?	HOW?
Business oriented		
Hands-on approach		
etc.		

- ✓ Explain the reasons of success: strong interaction with client? Deep expertise in AI?
- ➢ c. 1 slide

3

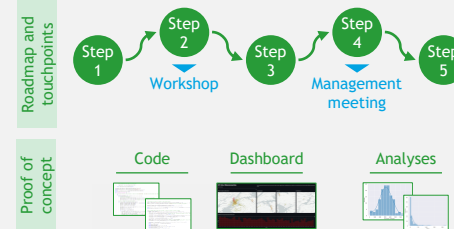Team: who are you and what have you done? Why are you relevant to solve the case?

- ✓ Quickly introduce yourself and why you are relevant for this mission
- ➢ c. 1 slide

4

Proposed methodology: what could you do for the potential client?

Roadmap and touchpoints

Step 1 → Step 2 (Workshop) → Step 3 → Step 4 (Management meeting) → Step 5

Proof of concept: Code, Dashboard, Analyses

- ✓ What you could do, how, what you have already done, any relevant insight to prove your expertise and how much time do you need for the mission!
- ➢ c. 6/8 slides

AGENDA

eleven augmented proposal challenge: deliver a convincing commercial exchange to your prospect leveraging your dual expertise in business and data science

Exercise:

✓ **Make groups** of 5 people

✓ **Choose** among the 2 available topics

✓ On your topic, **put yourself in a data consultant's shoes during a 1st commercial exchange**:

  ✓ What is the context of the company? (value chain, competition, technological trends, etc.)
  ✓ How could you generate added-value for them using potential of new technological enablers? (time-series, NLP, computer vision, etc.)
  ✓ How will the mission be structured? (roadmap, key steps, meeting with management, etc.)
  ✓ Can you make a first assessment of the feasibility of your solution? (data visualization, 1st demo, etc.)

✓ Note that the best solution for your client may not be the development of a custom-made solution, but could be the acquisition of a company, some key partnerships, etc.

✓ **Synthesize your work** in a commercial exchange (up to 15 slides)

✓ **Upload your work** on the Sharepoint (presentation, code, etc: see details at the end of this presentation)

✓ **Present to the jury** and prepare for questions

The Right Price

How to estimate a property price?

Airplane Interior Services
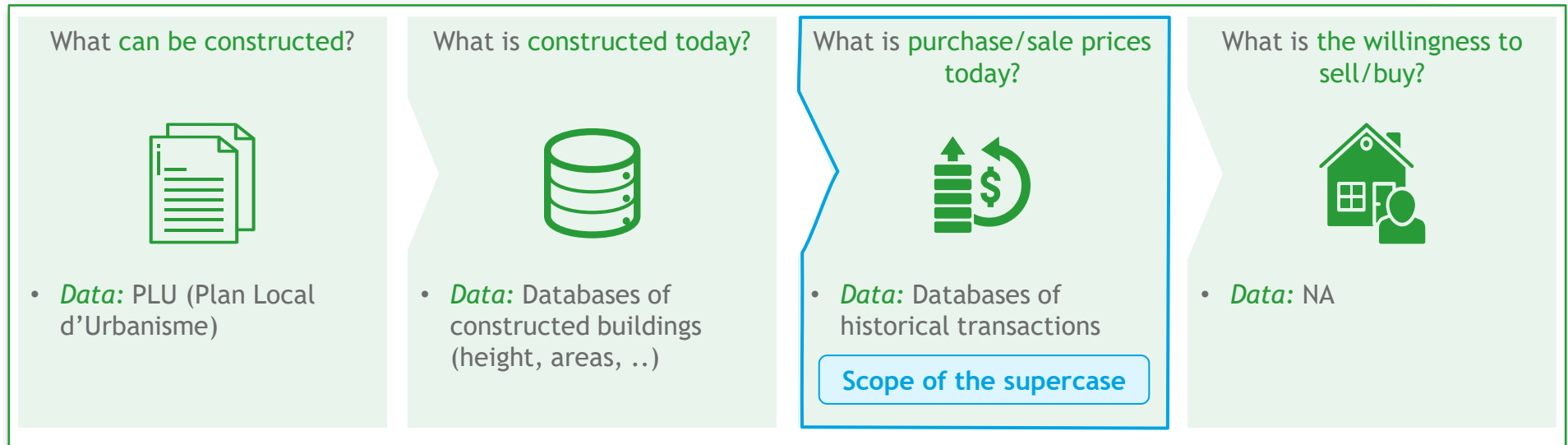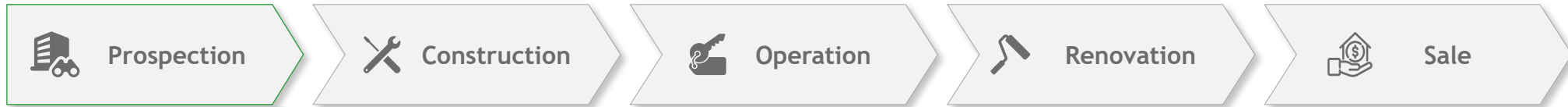
How to increase customers satisfaction?

The Right Price

improving definition and computation of mutability score in Île-de-France

Context: the potential client, a player in real-estate industry, wants to optimize its whole value chain, especially the purchase/sale price estimation

Value chain of a real estate development project

For illustration purposes

Prospection	Construction	Operation	Renovation	Sale

What can be constructed?

- *Data:* PLU (Plan Local d'Urbanisme)

What is constructed today?

- *Data:* Databases of constructed buildings (height, areas, ..)

What is purchase/sale prices today?

- *Data:* Databases of historical transactions

Scope of the supercase

What is the willingness to sell/buy?

- *Data:* NA

- Note that a price is useless without a related typology/area

E.g. In the same location a 60m² flat could be sold at 6000€/m² and a 40m² flat at 6500€/m²

Data available: you have access to geo-localized data from official mutation databases in Île-de-France, thereby providing valuable data sources with large and various information on the properties

Note that as in any data science project, the data may require further data engineering before being fully leverageable

datemut	anneemut	moismut	coddep	libnatmut	vefa	valeur fonciere	codtypebien	libtypbien	longitude	latitude	nbdispo	nblot	nbcomm	l_codinsee	nbsection
22/10/2018	2018	10	75	Vente	FALSE	147000.0	121	UN APPARTEMENT	2.3449403	48.860205	1	1	1	75101	1
26/10/2015	2015	10	75	Vente	FALSE	95000.0	131	UNE DEPENDANCE	2.3304056	48.866149	1	1	1	75101	1
13/06/2019	2019	6	75	Vente	FALSE	1900000.0	152	BATI MIXTE - LOGEMEI	2.3435634	48.861828	1	0	1	75101	1
20/04/2017	2017	4	75	Vente	FALSE	1509000.0	121	UN APPARTEMENT	2.3375203	48.865924	1	2	1	75101	1
06/03/2015	2015	3	75	Vente	FALSE	28250.0	131	UNE DEPENDANCE	2.3436309	48.860933	1	1	1	75101	1
01/08/2017	2017	8	75	Vente	FALSE	500000.0	121	UN APPARTEMENT	2.3427047	48.861218	1	1	1	75101	1
03/07/2019	2019	7	75	Vente	FALSE	367000.0	121	UN APPARTEMENT	2.3403178	48.863702	1	1	1	75101	1
27/10/2016	2016	10	75	Vente	FALSE	90000.0	121	UN APPARTEMENT	2.3487978	48.863630	1	2	1	75101	1
23/11/2018	2018	11	75	Vente	FALSE	135000.0	121	UN APPARTEMENT	2.3354345	48.866459	1	1	1	75101	1
26/08/2016	2016	8	75	Vente	FALSE	190000.0	121	UN APPARTEMENT	2.3416377	48.864750	1	1	1	75101	1
28/12/2017	2017	12	75	Vente	FALSE	300000.0	14	ACTIVITE	2.3473709	48.864122	1	2	1	75101	1
11/06/2014	2014	6	75	Vente	FALSE	730000.0	121	UN APPARTEMENT	2.3316604	48.864810	1	2	1	75101	1
02/05/2018	2018	5	75	Vente	FALSE	935000.0	121	UN APPARTEMENT	2.3293296	48.865974	1	1	1	75101	1
11/05/2016	2016	5	75	Vente	FALSE	535000.0	121	UN APPARTEMENT	2.3468805	48.859271	1	1	1	75101	1
07/10/2016	2016	10	75	Vente	FALSE	339050.0	121	UN APPARTEMENT	2.3361317	48.865615	1	1	1	75101	1
11/07/2018	2018	7	75	Vente	FALSE	413438.0	121	UN APPARTEMENT	2.3334257	48.864586	1	2	1	75101	1
22/10/2018	2018	10	75	Vente	FALSE	3327000.0	14	ACTIVITE	2.3407276	48.862311	1	3	1	75101	1
06/02/2018	2018	2	75	Vente	FALSE	650000.0	122	DEUX APPARTEMENTS	2.3331821	48.864667	2	2	1	75101	1
04/09/2018	2018	9	75	Vente	FALSE	375000.0	14	ACTIVITE	2.3452226	48.860565	1	1	1	75101	1
09/02/2015	2015	2	75	Vente	FALSE	163000.0	121	UN APPARTEMENT	2.3413226	48.865097	1	2	1	75101	1
18/06/2020	2020	6	75	Vente	FALSE	768490.0	121	UN APPARTEMENT	2.3437107	48.863758	1	1	1	75101	1
07/06/2017	2017	6	75	Vente	FALSE	15000.0	14	ACTIVITE	2.3428064	48.861055	1	1	1	75101	1
13/11/2014	2014	11	75	Vente	FALSE	480000.0	121	UN APPARTEMENT	2.3413863	48.861495	1	1	1	75101	1
30/03/2018	2018	3	75	Vente	FALSE	823000.0	121	UN APPARTEMENT	2.3330651	48.863992	1	2	1	75101	1
14/09/2015	2015	9	75	Vente	FALSE	250000.0	121	UN APPARTEMENT	2.3448141	48.858214	1	1	1	75101	1
16/02/2018	2018	2	75	Vente	FALSE	2873000.0	122	DEUX APPARTEMENTS	2.3354345	48.866459	2	3	1	75101	1
11/03/2019	2019	3	75	Vente	FALSE	525000.0	121	UN APPARTEMENT	2.3356954	48.865197	1	2	1	75101	1

Optional: Bonus point

✓ How would have you geo-localized this data by yourself?

Resources : You are free to use any resources you want, here are some recommendations to help you get started

Languages

We highly recommend you use Python even though same kind of results could be achieved with similar tools (with R for instance)
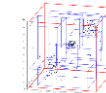
Relevant libraries

 scikit learn

 GeoPandas

 Rtree

 Streamlit

To develop a wide range of ML models

To use geo-localized visualization tools

To use geo-localized advanced ML models

To efficiently develop a dashboard / front-end

For information: the origin source of the data

| 75 Historical mutations | 77 Historical mutations | → | eleven secret recipe | → | 75 Geo-localized mutations | 77 Geo-localized mutations |
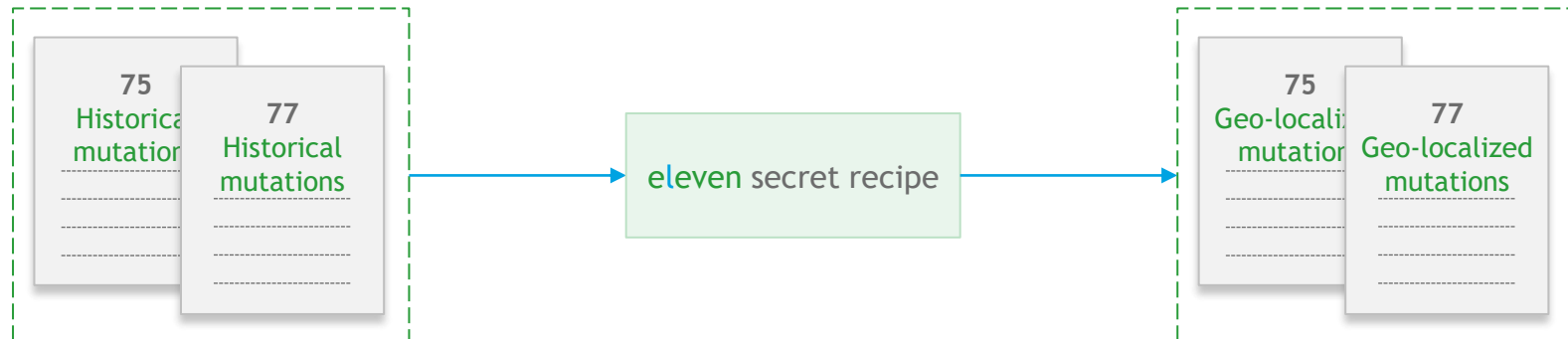
AGENDA

1. About eleven

2. Structure of a commercial exchange

3. Case presentations

   a) The Right Price

   b) Airplane Interior Service

4. General information

   a) Expected output

   b) Practical information

Airplane Interior Services

designing new airplane interior services based on user experiences

Context: the potential client, a player in the aeronautic industry, wants to understand how the user experience of airplane passengers can be improved

For illustration purposes

Traveler journey

Emotional curve and main pain points

Innovation opportunities

| Let's put PlayStations in the waiting areas | Let's change how the seats are assorted | Let's give more free food |

Scope of the supercase
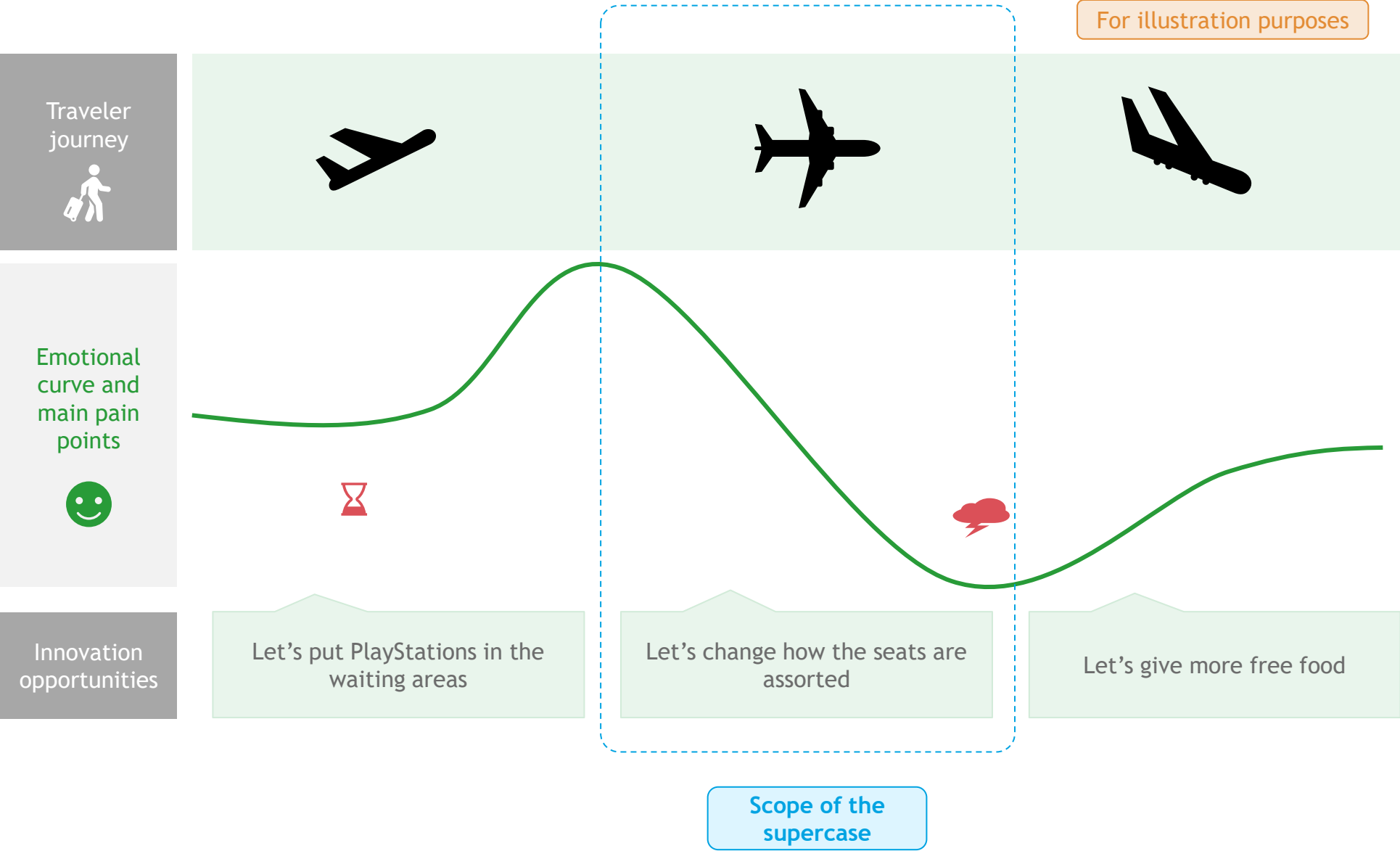
Data available: you have access to scrapped data from several websites enable users to share their experience, thereby providing valuable data sources with large and various information

Note that the data provided just comes out of scrapping process, hence it would require some data engineering before being leveraged

Optional: Bonus point

✓ How would have you collected this data yourself?
✓ Can you provide a quick demonstration of this working?

Resources: You are free to use any resources you want, here are some recommendations to help you get started

---

Languages

We highly recommend you use Python even though same kind of results could be achieved with similar tools (with R for instance)

Relevant libraries

To develop a wide range of ML models

To develop topic-modelling algorithm as well as sentiment-analysis tools

To efficiently develop a dashboard / front-end

For information: the origin source of the data

https://www.seatguru.com/browseairlines/browseairlines.php

https://www.seatguru.com/airlines/Aegean_Airlines/Aegean_Airlines_Airbus_A320-200.php

https://www.airlinequality.com/review-pages/a-z-airline-reviews/

Illustration: Topic modelling and sentiment analysis enable to derive valuable insights from customer reviews

NLP task

TOPIC MODELLING

Aims at extracting topics from text entities.

Typical approaches:
- Latent Dirichlet Allocation
- Classification
- Vectorization and clustering
- …

Typical challenges
- Number of topics
- Granularity of the analysis: full document, sentences, sub-sentences…

The service was bad to non-existent. 2 crew members for a rather large area.

SENTIMENT ANALYSIS

Aims at classifying a sentence as expressing a positive, negative or neutral opinion

Typical challenges
- Direct vs. comparative opinions
- Explicit vs. implicit opinions
- Granularity of the analysis: full document, sentences, sub-sentences…

Topic: crew

Sentiment: -0.8

AGENDA

1. About eleven

2. Structure of a commercial exchange

3. Case presentations

   a) The Right Price

   b) Airplane Interior Service

4. General information

   a) Expected output

   b) Practical information

Expected output: how should you structure a commercial exchange?

You are expected to deliver the following:

- Your code used to solve the use case, including any relevant dashboard, data analysis, model training, etc. This code should be readable by anyone, with a clear *Readme*, *requirements*, potential *quickstart.py*, etc.
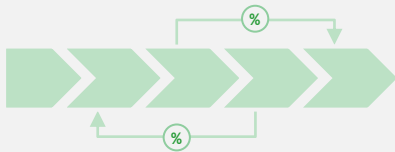
- A "client-ready" PowerPoint presentation of your work, structured as follows:

1

Context: what do you understand from the project? Why is the context very complex?

- ✓ Present your **understanding** of the situation
- ✓ Include the **process**, the **needs**, the potential **ROI**, etc.

  ➢ c. **2/3 slides**

2

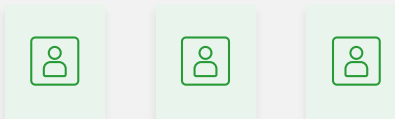**Key Success Factors**: what would be the reason of your success?

KSF	WHY?	HOW?
Business oriented		
Hands-on approach		
etc.		

- ✓ Explain the reasons of success: strong **interaction** with client? Deep **expertise** in AI?

  ➢ c. **1 slide**

3

Team: who are you and what have you done? Why are you relevant to solve the case?

- ✓ Quickly **introduce yourself** and **why you are relevant** for this mission

  ➢ c. **1 slide**

4

Proposed methodology: what could you do for the potential client?

Roadmap and touchpoints

Step 1 → Step 2 → Step 3 → Step 4 → Step 5
Workshop   Management meeting

Proof of concept

Code   Dashboard   Analyses

- ✓ What you **could do, how**, what you have **already done**, any **relevant insight** to prove your expertise and **how much time** do you need **for the mission**!

  ➢ c. **6/8 slides**

HEC PARIS

eleven strategy consultants

For illustration

For illustration

Based on past experiences, eleven has identified **five key success factors** to successfully launch a first AI initiative

KEY SUCCESS FACTORS	WHY?	HOW?
1 FOLLOW A BUSINESS-ORIENTED ADOPTION OF AI		
2 ONBOARD THE TEAMS AND BUSINESS STAKEHOLDERS EARLY		
3 MASTER AI TOOLS AND METHODOLOGY		
4 START SMALL THEN ITERATE		
5 FOLLOW A HANDS-ON APPROACH		

For illustration

eleven suggests to carry out the mission following **two phases** in order to **remove potential technological barriers** before launching the implementation phase
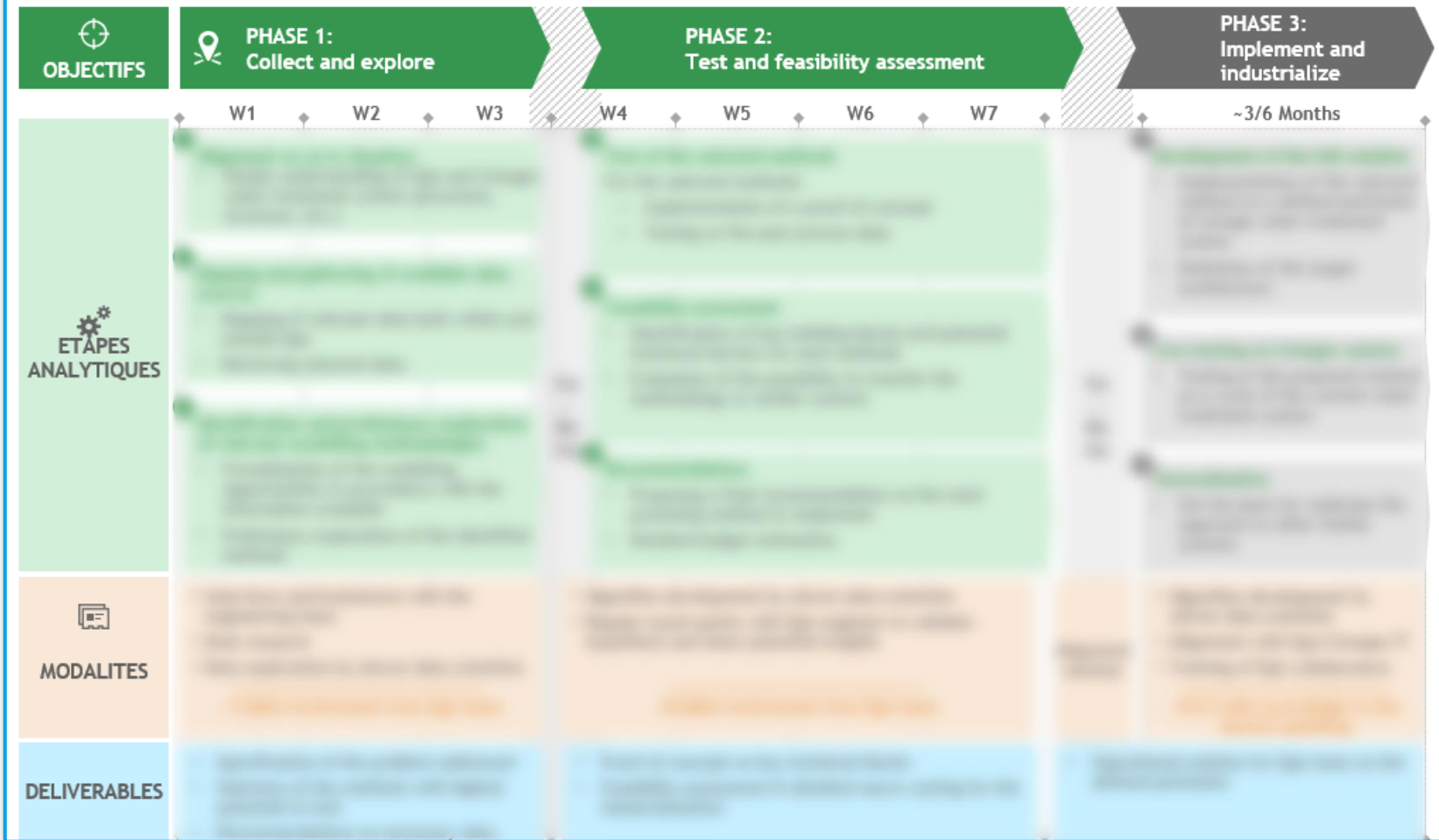
OBJECTIFS	PHASE 1: Collect and explore			PHASE 2: Test and feasibility assessment				PHASE 3: Implement and industrialize
	W1	W2	W3	W4	W5	W6	W7	~3/6 Months
ETAPES ANALYTIQUES								
MODALITES								
DELIVERABLES								

eleven
strategy consultants

AGENDA

1. About eleven

2. Structure of a commercial exchange

3. Case presentations

   a) The Right Price

   b) Airplane Interior Service

4. General information

   a) Expected output

   b) Practical information

eleven
strategy consultants

The Slack workplace

For this challenge, we opened a Slack workplace for you to ask your questions when eleven consultants are not on campus

On this workplace, you will find three channels:

1) **général**: for all questions and information related to the organization of the challenge

2) **TheRightPrice**: for all questions specifically related to *TheRightPrice* case

3) **AirplaneInteriorServices**: for all questions specifically related to *AirplaneInteriorServices* case

Additional information may also be pinned in these channels (schedules, classroom numbers, etc.)

Please use the right channel to ensure fluidity of the interactions

Before asking something, also make sure that the requested information has not been given already ;)

You may download the slack application on your device or access it via your usual browser

Download instructions & submission process

How to download datasets ?

You can download datasets and potential additional information at the following links:

- **Case #1**: AIS – get your files here

- **Case #2**: TRP – get your files here

How to submit your works ?

Each group will receive a link by email with a Sharepoint folder to submit their assignment (both Presentation + Code)

Notes:

- You can organize your folder as you wish.

- You can keep old files that should not be submitted to the jury in a *0_Archives* folder

Proposed schedule for the two days: the timing may be short, do not hesitate to split the work between the members of the team

For illustration

	Monday 31st	Tuesday 1st		Wednesday 2nd
AM session		**Business** ✓ Present your team and your relevance for the mission ✓ Describe the value chain of the client and the potential untapped levers		**Business** ✓ Evaluate the added-value your solution could provide to your potential client ✓ Complete the expected methodology to answer client's needs
		Data ✓ Start data exploration and cleaning ✓ Confirm analysis to be led or models to be used		**Data** ✓ Take a step back: evaluate the quality of your models ✓ Present your results – optional: presentation could include a dashboard
PM session	**Kick-off (*2pm - T201*)** **Working Time (T207 + T208)** ✓ Ensure your understanding of the value chain of the potential client ✓ Collect and explore the data ✓ Validate the analysis you would like to perform with the provided data ✓ Split the work between the team *Start sliding quickly, last minute slides could lead to poor final work quality*	**Business** ✓ Identify the Key Success Factors (strong interaction with client, hands-on, etc.) ✓ Start working on the methodology you could use to deliver a solution **Data** ✓ Perform any insightful analysis ✓ Build any relevant model	**Q&A (remote)** ✓ 30' per group ✓ Planning on slack	**Final Pitch (*1.30pm – T207 + T208*)** ✓ Be sure your slides are "client-ready" ✓ Upload both final presentation and final code ✓ Be prepared with all your speakers knowing their part **Closing Session (*5.30 pm - T201*)**

Final presentation details and best practices:

On Wednesday 2nd, you will have to present your work in front of a jury during a closed-door session

The modalities of the presentation will be as follow:

- 10min group pitch based on a PowerPoint presentation
- ~10min Q&A session with the jury
- ~10min debrief from the jury

For each case, a winner will be announced. The two winner groups will then present their work to the other students (same modalities with questions from the students)

The presentation must be as professional as possible. Here are some advices and best practices that may be useful:

- **Structure your presentation**: start by stating the problem that you want to solve, then present the way you tackled it, and finally describe your solution. The "story" of the presentation should be natural and easy to follow

- **Be concise and precise**: focus on the most important messages, as you only have 10 minutes to present the work achieved. You should limit the number of slides you present (you can still add appendices if needed)

- **Be organized as a team**: split up the speaking time between the team members beforehand to make it smoother

- **Be honest**: tell where you encountered issues or challenges

- **C-suite level**: you should convince both the CEO and the CTO/CDO of the company

Evaluation criteria:

Although different in their essence, the cases will be graded based on similar criteria.
NB: any provided code will be tested in order to confirm its good functioning.

	Topic	Description	Contribution to the final grade*
	Engagement	The engagement of the team during the exercise (how far you've gone, how autonomous you have been, how much you have asked questions when stuck, etc.)	2 points
Business aspect	Presentation quality	The quality of your final presentation: how professional it looks (slide quality), how clear and complete it is (storytelling), how pertinent your answers are, etc.	3 points
	Business methodology	The creativity and relevance of the methodology (i.e. scientific approach) you choose regarding the problem you try to solve, and the data provided, the business sense behind your methodology and the pragmatism of your presentation	6 points
Technical aspect	Technical choices	The explanation of your technical choices and your ability to present them in non-technical terms	3 points
	Model efficiency or Analysis relevance	The performance of your model (specific to each case), the relevance for the problem in question, the quality of the analysis led	5 points
	Code good practices	Your code must be well structured, easy to run and easy to understand with clear readme and requirements.	1 point

Please note that all groups will be graded at the end of the week

Enjoy the challenge!