Airbnb

Airbnb Business Analysis Track 1

University of Essex Online

Module: Machine Learning

Group D

Second group meeting 9.5.2026

Wrap-up

Ariel Mella <ariel.mella@gmail.com> Fri, May 8, 10:25 AM (1 day ago)
to Sarra, Francesca, Lyth, Tamim, Hadeel

Hello all,

Yesterday we had our first group meeting, namely the kick-off. many thanks and a really constructive interlock.

We have discussed, agreed, and made some decisions in order to move forward.

From now on, all meeting minutes and any assets from the regular meetings will be saved in the Meeting Minutes folder.

General Google Drive for Group D link:
Link: https://drive.google.com/drive/folders/1vcAAdPquTZ693DTfI5C46Oqau9pwMips?usp=sharing

Meeting Minutes folder:
Link: https://drive.google.com/drive/folders/11b7-MjSMYRwSH_TF7MUE9_ySXrpGZehc?usp=drive_link

The presentation that we used for our kickoff discussion was shared on a previous email and is within the drive you should all have access already to it.

The drive will get populated with more folder structure as the next activities develop.

Key next steps:

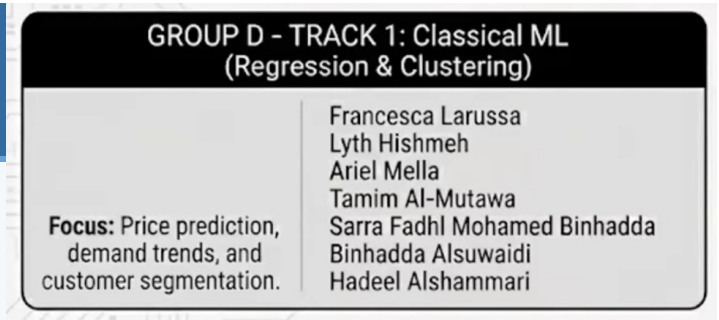
1. Next meeting on Saturday, May 9th, at 1 PM UK time. Invites have been sent to all.
2. In case Lyth can't join us on the next call, Ariel has separately sent a proposal on design decisions related to the business question, workstreams, and roles. We would really welcome Lyth's feedback, if available, as it would help us move forward with the specific task allocation.
2.1 We of course welcome Tamim or Hadeel if you could join us, we will catchup and understand what is the best task allocation we can do collectively.
3. For anyone who could not join the kick-off, please go through the meeting minutes in detail to understand what was discussed and agreed, and what comes next. Also, please go through the presentation before joining the next call, if you have not done so already, so everyone is up to speed with the proposals. We would really welcome your comments, feedback, and counter-proposals, especially on the business question, workstreams, and roles, so we can have a productive discussion and move forward together.

Security note: As not everyone has a Google account, the sharing was done via a mix of Google permissions and "anyone with the link can view." Therefore, please do not share the links with anyone outside Group D.

Many thanks and looking forward to the task allocation so we can get hands-on!

Ariel

Team Status and Task allocations



1) As we could not talk with Lyth, Tamim and Hadeel we miss the brief about them (about yourself, skills and interests related to this assignment, availability to work and communicate, target grade and objective)
2) Tamim agreed that today we take the decisions (workstreams, business question, roles and distribution) without him.
3) Ariel offered via Whatsapp to Tamim the role of QA/reviewer of the final written report, to be discussed today on Tamim's absence.
4) As there was no deep interlock with the rest of the team members the proposal is to move forward the ones present in the call with all decicions and task allocations.

Proposed business question

Task 2 (Data Analysis – Based on Track Selection)

Each group must apply their assigned machine learning approach to analyse the dataset.

Track-Specific Expectations:

- Track 1 (Classical ML – Regression & Clustering): Apply regression models for pricing prediction and clustering techniques for customer segmentation.

Task 1 (Business Analytic Question)

Each group must propose a business-oriented question relevant to Airbnb. The question should be practical and valuable to Airbnb's strategic planning. An example might be: "How do seasonal fluctuations affect Airbnb pricing strategies in different neighbourhoods?"

Airbnb dataset as per assignment.

Regression for price prediction.

"Can Airbnb identify pricing patterns and listing segments in the NYC market to support hosts with competitive pricing and customise segment-specific guidance?"

Clustering segmentation.

Business value supporting hosts.

Rational and business outcome: by applying classical ML, we can deliver key information to the business for better decisions: EDA should unearth demand patterns and indicators from reviews and availability, regression will support right pricing decisions, and clustering supports segment grouping so Airbnb and hosts can potentially steer or increase pricing levels without negative demand impact and/or adjust pricing to support demand.

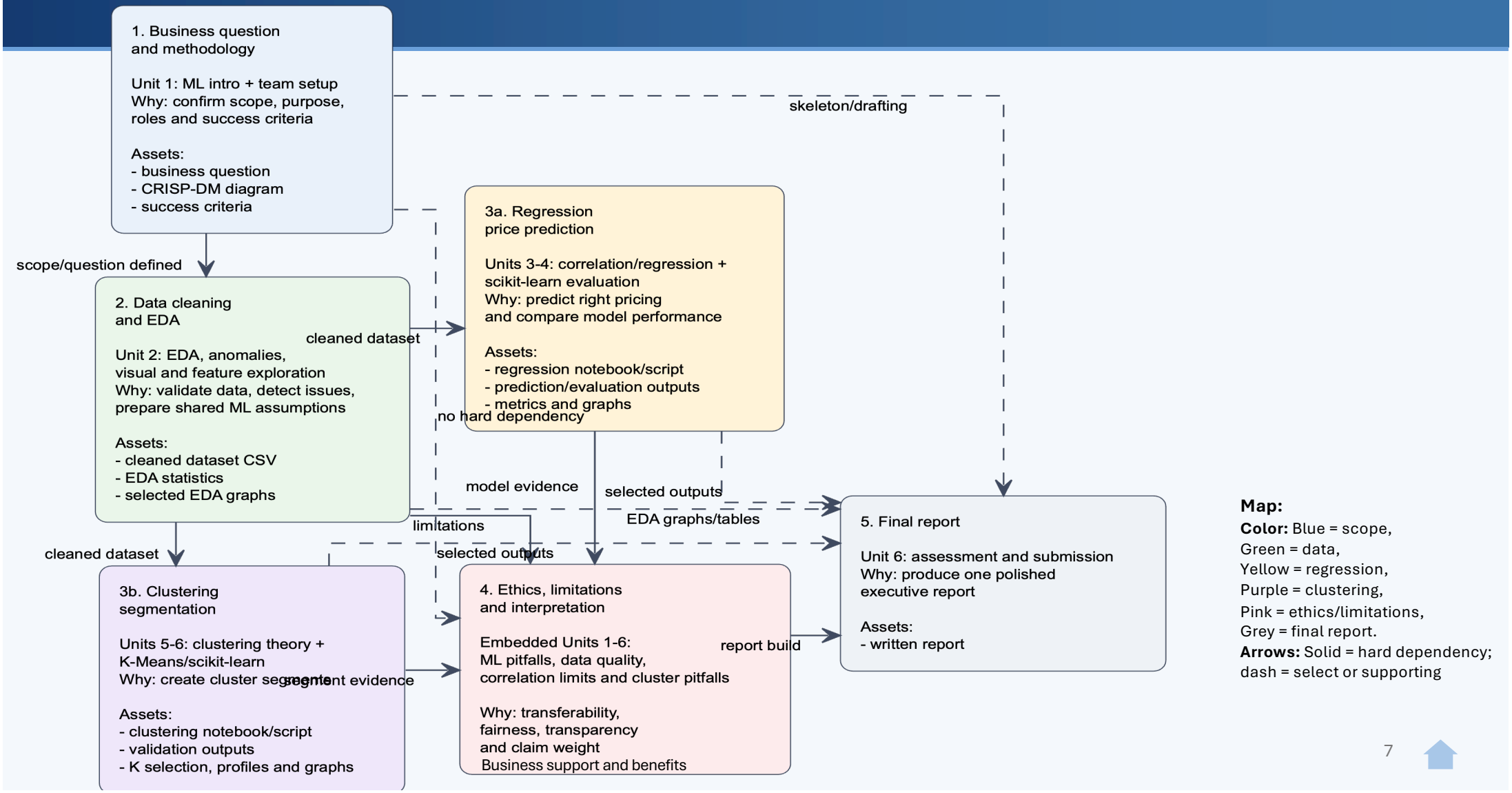
Activities high-level

- **Data cleaning:** prepare dataset, define how to deal with missingness, 0 values, outliers and proxy variables. Thus, both ML workstreams use the same consistent dataset.
- **EDA:** should unearth the demand patterns and indicators by reviewing price, room type, location, reviews and availability, using statistics and graphs that directly support the business question.
- **Regression for price prediction**: predict right pricing, compare model performance and produce evidence that can support Airbnb and host pricing decisions.
- **Clustering grouping**: K-Means to support segment grouping, creating cluster segments that help Airbnb understand which listing groups may need different pricing, demand and host-support actions.
- **Writing the report:** from skeleton to draft, to final submission. Can be fed with different inputs alongside the journey, and obviously with all relevant sections.
- **QA/ Review:** final review and quality assurance, content wise (classical ML applied, EDA, business factors) and academic writing wise.
- **Coordination:** general coordination and maintenance tasks, tooling, facilitate that tasks are done on time with the expected quality and everyone's have what they need. General communication.

Proposed roles, task allocation and ownership (roles, not # people)

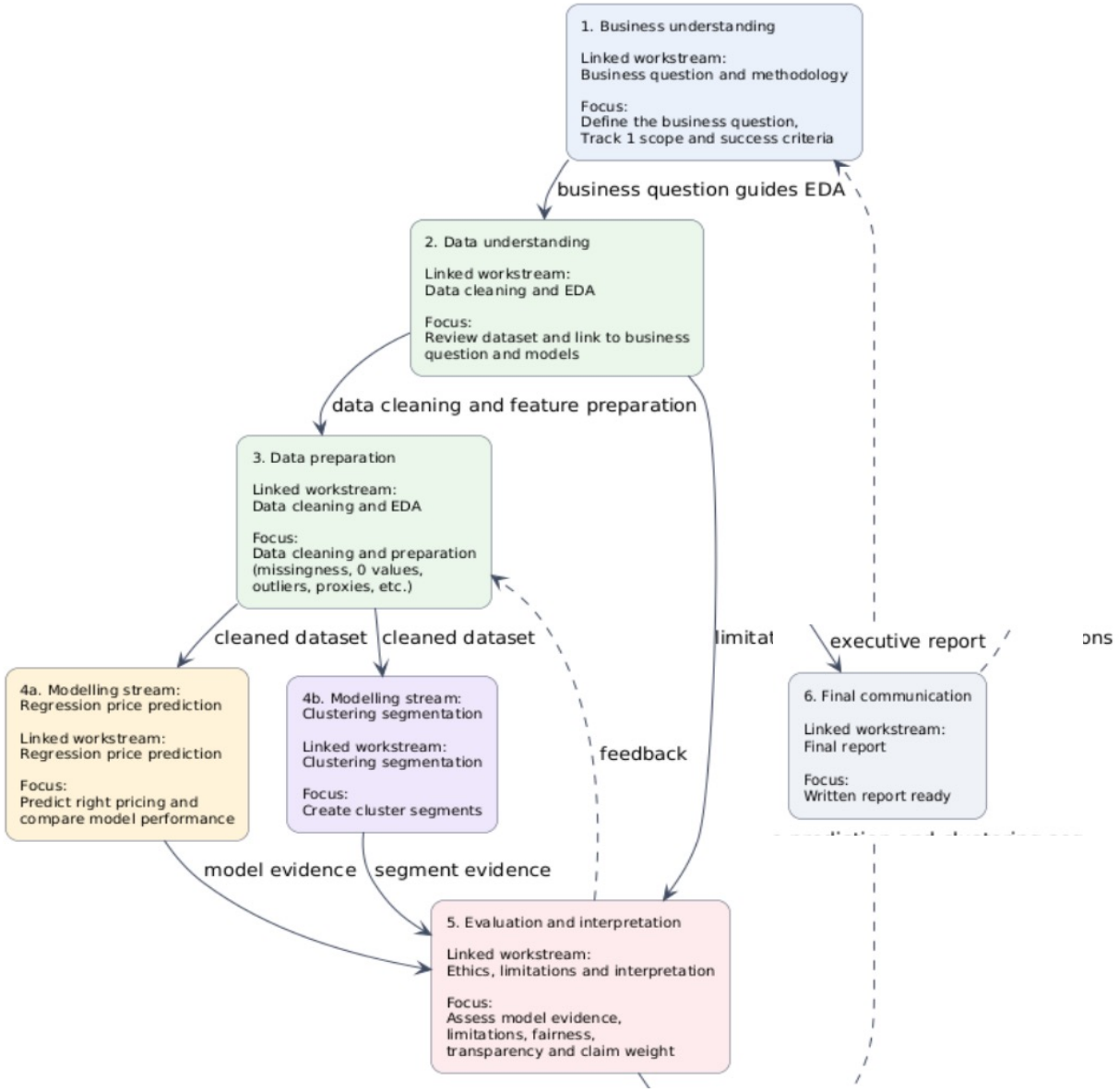
Role	Accountable for	Outcome	Best fit / skills (actual and to learn)
Report Owner and Integrator	Full written report from skeleton to final version	FINAL version	Academic writing, business, general ML, research
Data Cleaning + EDA	Dataset cleaning, EDA statistics and graphs	Cleaned dataset, EDA outputs, key patterns	Python/Pandas, statistics, data quality
Regression	Price prediction modelling and evaluation	Python Regression code, metrics, model evidence, graphs	Regression, scikit-learn, evaluation
Clustering	K-Means segmentation and validation	Python Clustering code, K selection, segment outputs, graphs	K-Means, validation, segmentation logic
Ethics + Limitations	Risks, limitations, claim weight, references, criticality	Ethics notes, limitations, track comparison	Critical thinking, ethics angle
QA/Review	Final review and quality assurance	Final document ready to submit	Academic writing, classical ML, business, research, critical thinking
Coordinator	Task coordination and facilitating, communications and admin stuff,	Basic project admin tasks Tasks on time	Coordination, project management

Group D - Track 1 WBS: unit links, dependencies and assets



Map:
Color: Blue = scope,
Green = data,
Yellow = regression,
Purple = clustering,
Pink = ethics/limitations,
Grey = final report.
Arrows: Solid = hard dependency;
dash = select or supporting

Proposed approach using CRISP-DM like



Timeline TO BE ADJUSTED

Week / dates	Task	Owner	Input	Output
By 16/May	EDA + Data Cleaning	Francesca	Ariel provides basic design and plots/charts ideas	Business question, success criteria, report skeleton, CRISP-DM view
W2 5–11/May	Data cleaning and EDA	Data Cleaning + EDA Lead	dataset, business question	Cleaned dataset CSV, EDA statistics, selected EDA graphs
W3 12–18/May	Regression price prediction	Regression Lead	Cleaned dataset, EDA signals, pricing target	Python notebook/script, prediction/evaluation outputs, metrics and graphs
W3 12–18/May	Clustering segmentation	Clustering Lead	Cleaned dataset, EDA signals, selected variables	Python notebook/script, segmentation/validation outputs, K selection, profiles and graphs
W4 19–25/May	Ethics, limitations and interpretation	Report Owner / Integrator or Ethics + Limitations Lead	EDA limitations, model evidence, segment evidence	Claim weight, fairness/transparency, responsible use, what we can/cannot claim
W5 26/May–1/Jun	Final report build	Report Owner / Integrator + all leads	Selected outputs, handover notes, graphs/tables, references	Written report draft ready
W6 2–8/Jun	Submission	Report Owner + all QA	Full and pre final report, feedback, graphs/tables, references	Written report reviewed by everyone by ??/Jun; official deadline 8/Jun. submission.

Questions ?

Wrap-up from today
Next: who does what, by when ?

- Documented Key agreements and decisions, see details in:
- 1) Meeting notes shared via email,
- 2) replay the recording shared via email and
- 3) Definitions and next steps.9.5.2026 file containing key high level definitions including task allocation and key dates

Link: https://docs.google.com/presentation/d/1UiWfIjsBltygNyDvYrIfy-vQ2bOzGr2Z/edit?usp=drive_link&ouid=113786110083556405902&rtpof=true&sd=true

- Next slide: email with all relevant links (duplicated, was already sent via email). Note: everyone has access to all files regardless of the links, pleas go to the Google Drive root folder where all assets and working documents resides.

Meeting summary, decisions, definitions and document location email

Ariel Mella <ariel.mella@gmail.com> Sat, May 9, 4:32 PM (
to Sarra, Francesca, Lyth, Tamim, Hadeel

Hello all,

We had our group meeting today, and I'm pleased to share the key progress in our collective activities.

A high-level summary of the milestones is below, and you can find the details in the different links:

Presentation with all content reviewed today:
https://docs.google.com/presentation/d/1vRu0UP4VILUCHTppSbUComUBh8-hIbCL/edit?usp=drive_link&ouid=113786110083556405902&rtpof=true&sd=true

1. We have finalized the business question definition; details are in the link below.
2. We agreed on a timeline until completion; details are in the link below.
3. We agreed on the task distribution; details are in the link below.

Having decided on the key definitions needed we can now move to get hands-on, which is a great news.

For details please I do encourage each of you to go through below:

Meeting notes, summary and de-brief:
https://drive.google.com/drive/folders/11b7-MjSMYRwSH_TF7MUE9_ySXrpGZehc?usp=drive_link

Timeline and task allocation:
https://drive.google.com/drive/folders/1hPhM_Kf3iHUiPNMS3oJmbmHVPPz3LSj3?usp=drive_link

Business Question:
https://docs.google.com/presentation/d/1yz0DdIx9lntPOut7er8X1A-5HYW-YwEI/edit?usp=drive_link&ouid=113786110083556405902&rtpof=true&sd=true

We have agreed that the work now will be managed between the responsible people on 1:1 interlocks, so there is no need to have a group meeting next week.

Our next group meeting is on Saturday, 23rd May, which is already in our calendars.

Many thanks for your contribution, and looking forward to seeing great work come out of it.