



## Technical Assignment

Welcome! This practical assignment mirrors a real demand-planning problem. You'll work with two small datasets and a utility module, then present your approach and findings.

**Business goal:** Predict the coming week's demand 8 weeks in the future for a product at a supermarket and explain how your approach can reduce stock-outs and write-offs.

---

### What you receive (in the Case zip)

```
1 case/
2 └── demand.csv
3 └── promotions.csv
4 └── utils.py
```

You will create your own notebooks or scripts (e.g., `eda.ipynb` / `eda.py`, `model.ipynb` / `model.py`). A Git repo is optional and working locally is fine.

---

### What you need to deliver (how to submit)

Please email back a single ZIP (or a link to your own repo if you prefer) containing:

1. PR exercise

- Your docstring added to `clean_demand_per_group` in `utils.py`.
- A short PR review of `utils.py` (e.g., `PR_COMMENTS.md`).
- (*Optional*) Notes on better tests for `merge` or any unexpected behaviour you found.

2. EDA

- Your own notebook or script: `eda.ipynb` (or) `eda.py`.

3. Modelling

- Your own notebook or script: `model.ipynb` (or) `model.py`.

4. One-slide executive summary

- `conclusion_slide.pdf` (PDF or PPTX).

5. How to run

- `README_RUN.md` with clear steps to reproduce (environment + commands) and your time spent.

**Deadline:** Please submit 1 day before your interview so reviewers can prepare.

---

### Time guidance (indicative)

- ~30 min PR exercise
- ~60 min EDA
- ~60 min Modelling
- ~15 min One-slide conclusion

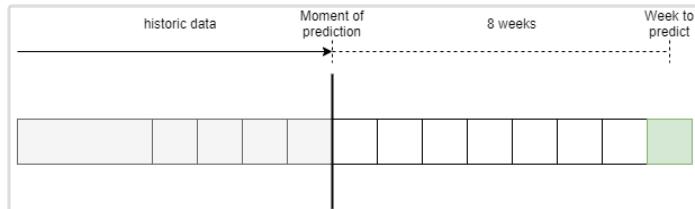
Spending more time is fine but not required. We assess in the context of the time you report.

## Business context

Our Global Analytics team supports data challenges across the organization. A colleague in Demand Planning (Netherlands) raised concerns:

“After our senior demand planner retired, we've seen more stock-outs for some products and more write-offs for others. Our forecasting has slipped; we need help.”

**In short:** Predict the demand for a product eight weeks ahead (the demand for the week that is 8 weeks in the future).



## Data description

- `demand.csv` daily demand for one supermarket × one SKU  
Columns: `date` (YYYY-MM-DD), `demand_units` (int  $\geq 0$ )
- `promotions.csv` promotion start dates (each promo lasts 7 days)  
Columns: `promo_start_date` (YYYY-MM-DD)

You may engineer features (e.g., weekly aggregates, calendar flags, lags, rolling stats, promo windows). State any assumptions.

## 1) Pull Request exercise (communication & code comprehension)

This task assesses how you read code, communicate feedback, and propose practical improvements. We recommend doing this first as `utils.py` can support your EDA/Modelling.

### Your tasks

1. Add a docstring for `utils.py::clean_demand_per_group` (use the NumPy style shown below).
2. Provide review comments on `utils.py` in a short `PR_COMMENTS.md` (prioritize correctness, clarity, robustness, tests).
3. (Optional) Suggest better tests for `utils.py::merge` or note any unexpected behaviour you found (with a minimal repro).

### Docstring style (example to follow)

```
1 """
2     Calculate the volume in hectoliters (vol hl) per product in a
3     DataFrame.
4
5     This function applies various extraction and conversion functions
6     to calculate the vol hl for each product
7     based on the 'name' column in the DataFrame.
```

```

7     Parameters
8     -----
9     df : pd.DataFrame
10    The DataFrame containing product details.
11
12    Returns
13    -----
14    pd.DataFrame
15    The DataFrame with calculated vol hl per product.
16
17    """

```

## 2) Exploratory Data Analysis (EDA)

Goal: assess data quality, understand time-series characteristics, and propose features relevant to an 8-week-ahead forecast.

Do at minimum:

- Data quality
- Patterns (trend/seasonality)
- Feature ideas
- Keep the narrative focused on what informs the 8-week horizon.

## 3) Modelling (forecasting 8 weeks ahead)

We're not chasing heavy optimisation—show sound judgement and correct evaluation.

Do at minimum:

1. Target definition: forecast the weekly demand 8 weeks ahead. (You may aggregate daily → weekly; please justify your choice.)
2. Baselines: include at least one simple baseline
3. Model: choose a reasonable method for the data volume
4. Evaluation: use time-aware backtesting with an 8-week horizon; report MAE and WAPE (add sMAPE/RMSE if useful).
5. Promotions: incorporate or justify their treatment
6. Explainability & operations: briefly show what drives predictions (if applicable) and outline how you'd deploy/monitor.

## 4) One-slide executive summary

Prepare one slide that answers:

“Can we predict 8 weeks ahead well enough to reduce stock-outs and write-offs?”

Include one clear chart/table, 1–2 key drivers (e.g., promos), and 2–3 bullet next steps. Keep it business-friendly.

### Interview format (≈1 hour)

1. Your one-slide (5–7 min)
2. Walkthrough of your approach (EDA → Modelling → PR highlights) (~30 min)
3. Q&A with two data scientists (10–12 min)
4. Your questions (~5 min)

## 5. Next steps (1–2 min)

---

### Environment & reproducibility

- Python 3.9+ recommended.
  - Suggested libs: `pandas`, `numpy`, `scikit-learn`, `matplotlib/plotly`, (optional) `statsmodels`, `xgboost`/`lightgbm`, `prophet`, `pmdarima`.
  - Include a `requirements.txt` (or `pyproject.toml`).
  - Add a seed where relevant; keep notebooks clean and runnable.
- 

### Assessment (how we'll look at your work)

- PR review quality (docstring clarity, actionable comments, test thinking)
- EDA depth (data quality, insightfulness, features tied to the 8-week horizon)
- Modelling (baseline, method appropriateness, leakage control, correct backtesting, metrics)
- Communication (structure, clarity, one-slide)

We consider your reported time spent and value clear trade-offs over excessive complexity.

---

### Questions or issues?

If anything blocks your progress, please reach out: [bryce.senekal@heineken.com](mailto:bryce.senekal@heineken.com),  
[jessica.matthysen@heineken.com](mailto:jessica.matthysen@heineken.com)

Good luck—we look forward to your approach!