# Suggested Order Quantities

MBAn Capstone project
AB InBev, Summer 2019, NYC


**AB InBev manager**: Bruno Raposo
**Faculty advisor**: Negin Golrezaei
**PhD mentor**: Andreea Georgescu


Antoine Roncoroni


Alexandre Saillard


## Problem Statement

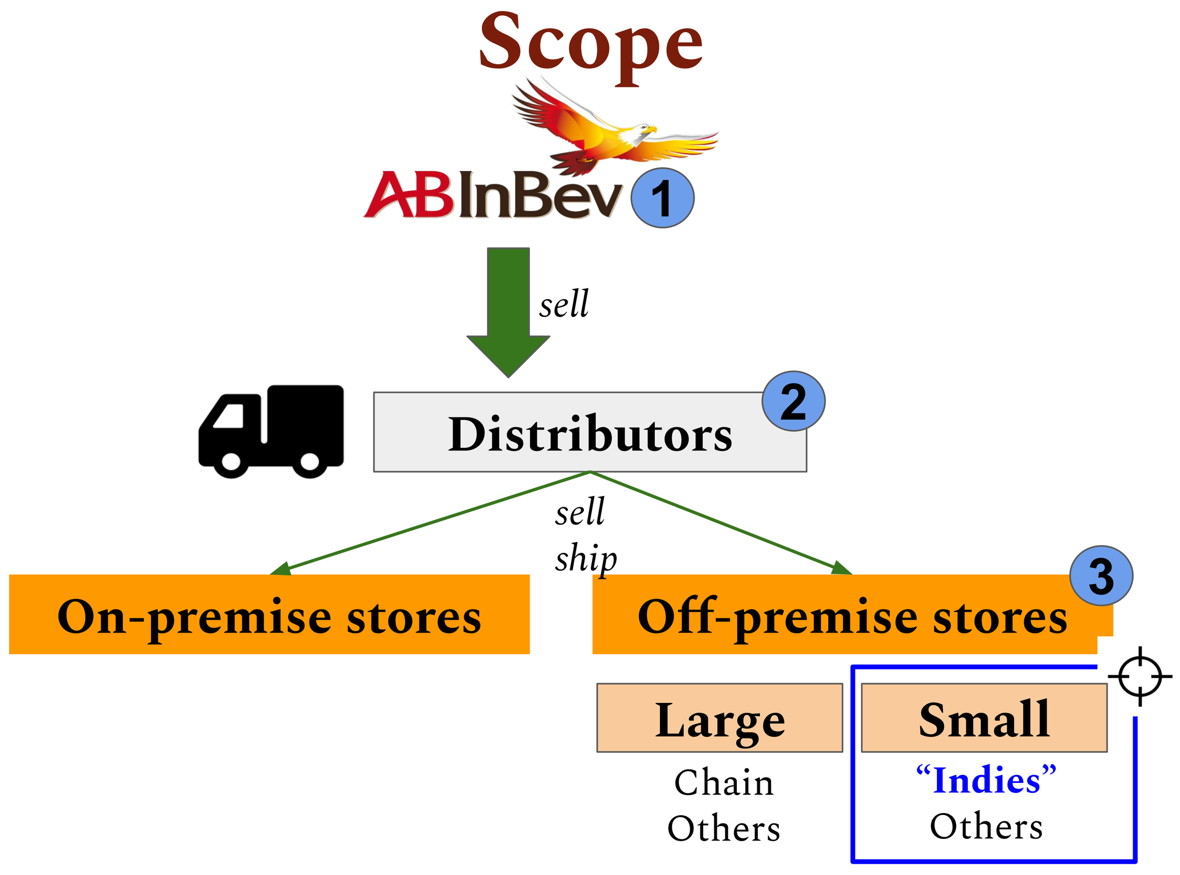
- US small independent stores brought **more than 20% of AB InBev's net revenue** in 2018
- Order placement involves a sales rep who spends time **counting inventory and visiting the store**
- Suggesting order quantities can unlock cost savings by reducing physical visits as well as better exploiting business opportunities e.g. upselling or assortment considerations **in a ~$5B market**

**Goal** — Predict

| Store | Date | Product | Quantity |
|---|---|---|---|
| X | 07/01/19 | BUD324 | 2 |
| X | 07/01/19 | BDL025 | 0 |
| X | 07/01/19 | STA034 | 1 |

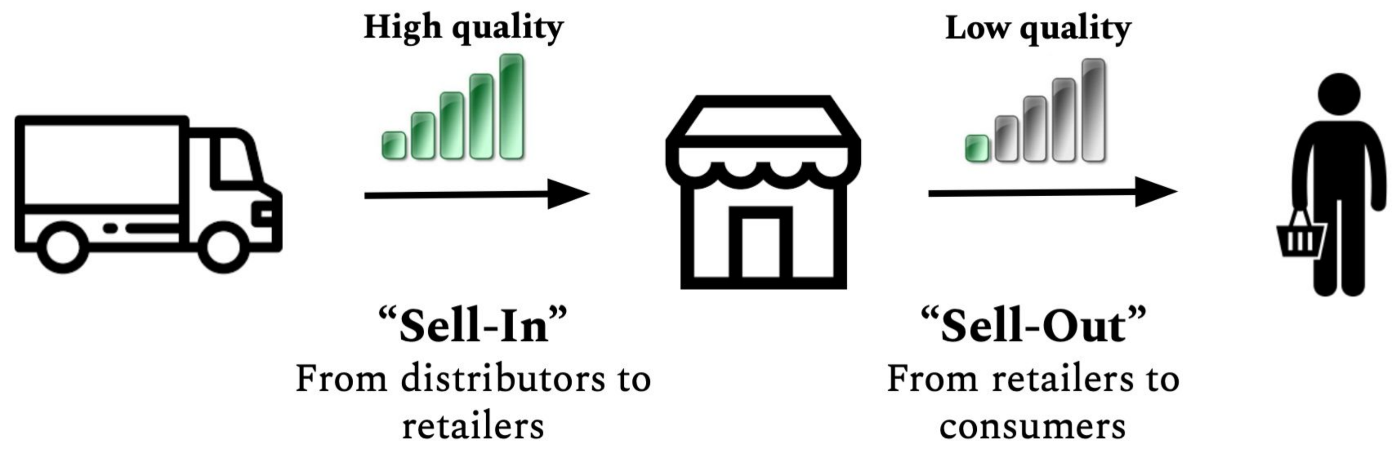
Project Kickoff 06/12 — Data Acquisition 06/21 — Field trip 06/25 — *Processing* — Mid-summer presentation 07/12 — *Modeling* — Problem re-scaling 08/09 — MIT Final presentation 08/23

## Scope




Scope: 60 small "indy" stores in the New York area

## Data




### Main data challenges

- **High granularity and noise** - Store-SKU-week level
- **Critical data availability** - No inventory data
- **Partial information** - only a fraction of sell-out recorded

60 Fraction (%) of sell-in recovered

- **Tracking of products**

"Matching issue": tracking products from the distributor (sell-in) to the end-consumer (sell-out)

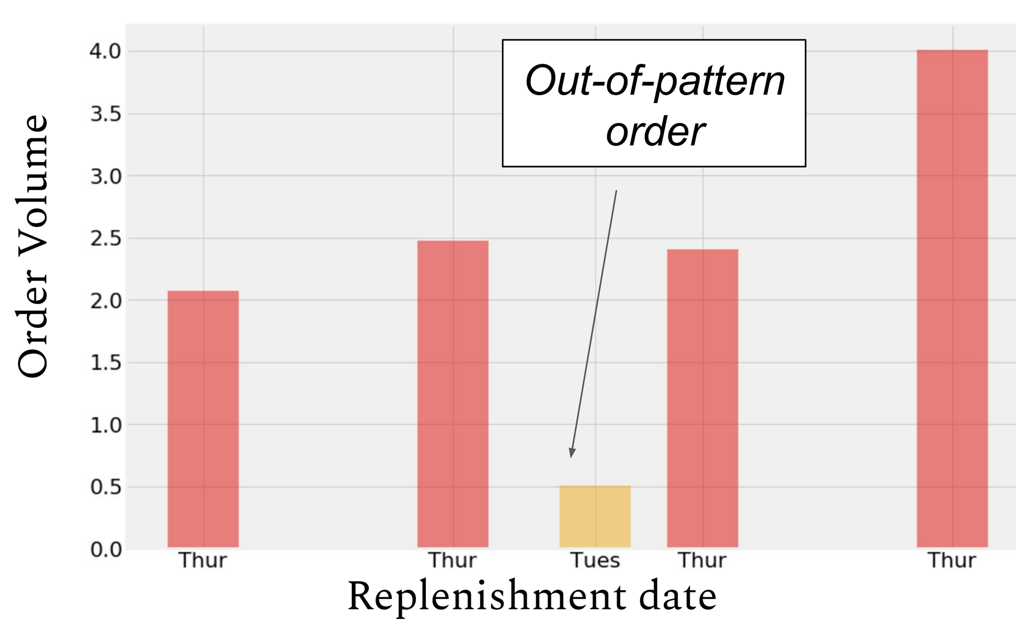
*SKU BUD004* — Unit sales — *single BUD 12oz bottle*

## Data Processing & Insights

Main challenges:

- **High # of historical SKUs**
  Some SKUs are hardly ever ordered, while some almost always
- **Erratic replenishment patterns**
  Order frequency is not fixed. Stores have different sizes and varying replenishment patterns.
- **Many gross outliers**
  Some orders do not fit into the store's replenishment pattern (ex: small order after an unexpected stock-out situation)

Evolution of order volume (store 'NU669')

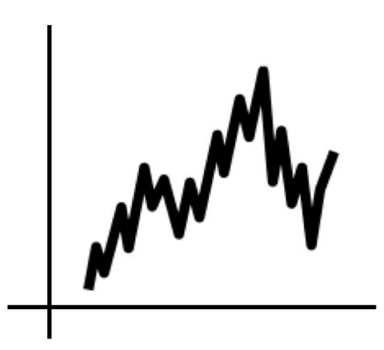



*Core set of SKUs*
Set of SKUs that account for 90% of historical sell-in sales

*Clustering of stores*
Cluster of "good" stores with weekly replenishment patterns

*Filtering of orders*
Gross outliers removed (small out-of-pattern orders)

## Modeling

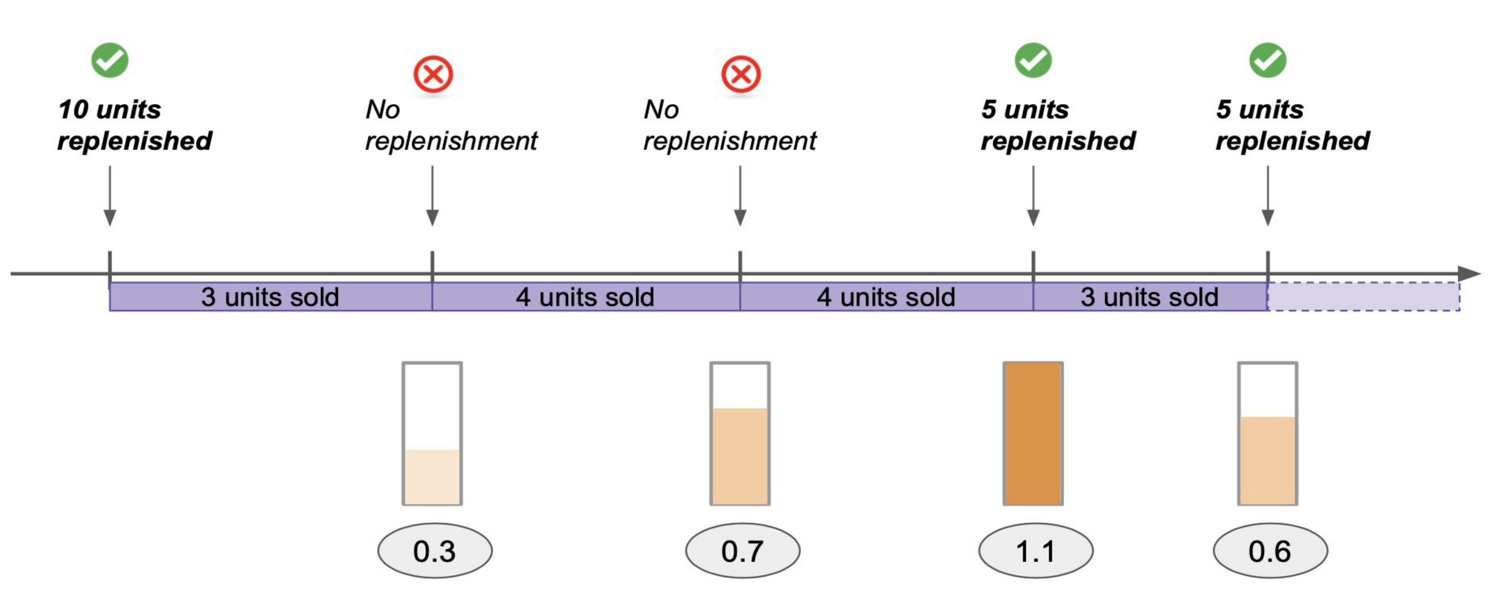


**Time series approach** (ARIMA/LSTM type models) — Model selection — **Micro approach** (consider exact sales/inventory quantities)

Given the extreme granularity of the task, a **feature-based approach** mimics the sales rep's decision process and thus seems the most appropriate.
Models implemented: Linear Regression, Random Forest
Feature selection: LASSO and forest feature importance

### Selected features

- **Recent replenished quantities**
  Rolling mean of # units replenished when actually ordered with or without including null values
- **Recent replenishment timing patterns**
  # consecutive replenishments without including a given SKU
  Normalized lead time
- **Sales trends and inventory status**
  Last period sell-out normalized by rolling mean of sell-out
  Stock consumption indicator
- **Next week's sales information**
  Weather forecast
  Binary incoming special event variable




**Stock-consumption indicator**
What fraction of my inventory did I consume since last replenishment?

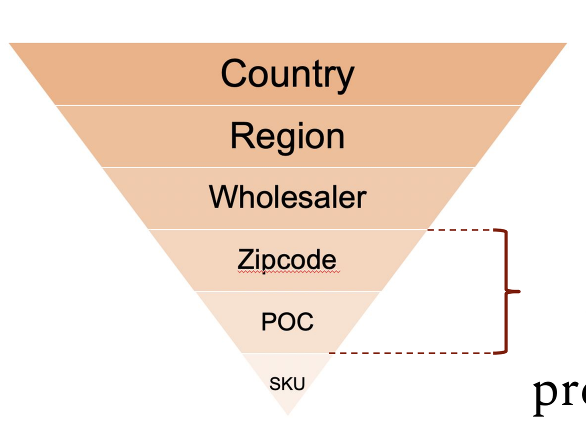
## Impact & results

Linear Regression trained at the store/SKU level for 8 SKUs with good matching results between sell-in and sell-out

| | Metric | Baseline | Linear Regression |
|---|---|---|---|
| Order basket | True Positive Rate | 66% | **72%** |
| | False Positive Rate | 17% | **16%** |
| | Precision | 59% | **63%** |
| Quantities | In-Margin Rate *on True Positives* | 77% | **79%** |
| | RMSE (% of mean) *on True Positives* | 43% | **41%** |

### Main takeaways

- Extreme granularity of task requires **accurate sales and inventory information**, enabling an optimization approach towards profit maximization.
- Acquisition of POS data is **costly** and remains **approximate**.
- Prediction task could focus on a more aggregated level, leaning towards **recommendations at the store-SKU level**.




Volume-level prediction e.g. bbl → Assortment reco

⇒ Recommendation broken down at the SKU level