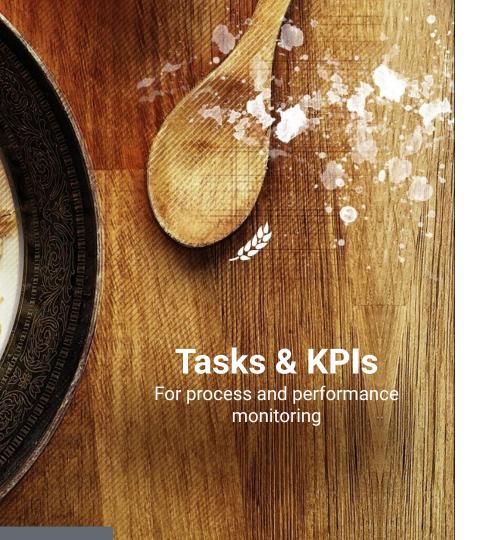
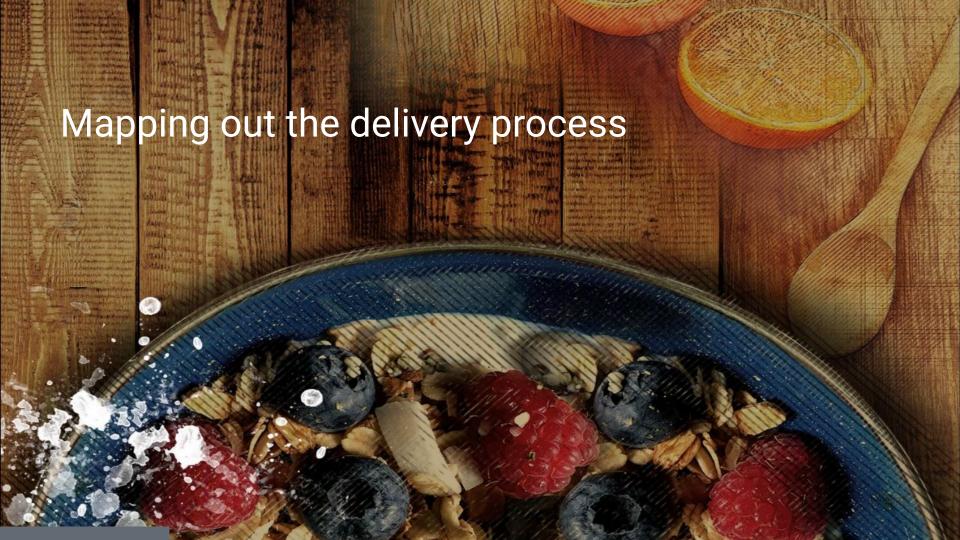




- Company provides 4 datasets
 - Orders
 - Order Process Data
 - Campaign Data
 - Intern Study Data
- Period 03. January 2017 30. December 2020 for orders only!
- Company deems having full visibility on order date and on truck date
- Other datasets sporadically cover orders in 2019 and 2020



- 1. Review & visualise process
- 2. Process time per order
- 3. Process time per ship mode
- 4. Track delivery time
- 5. Proof process assumptions
- 6. New KPIs & Improvements



Order in





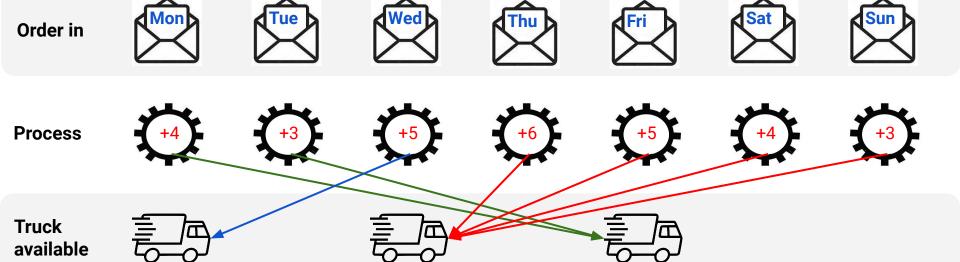


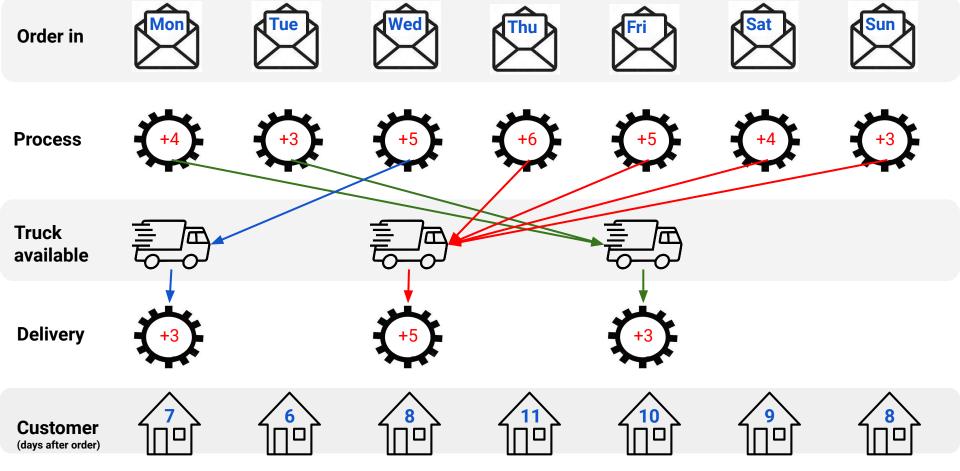


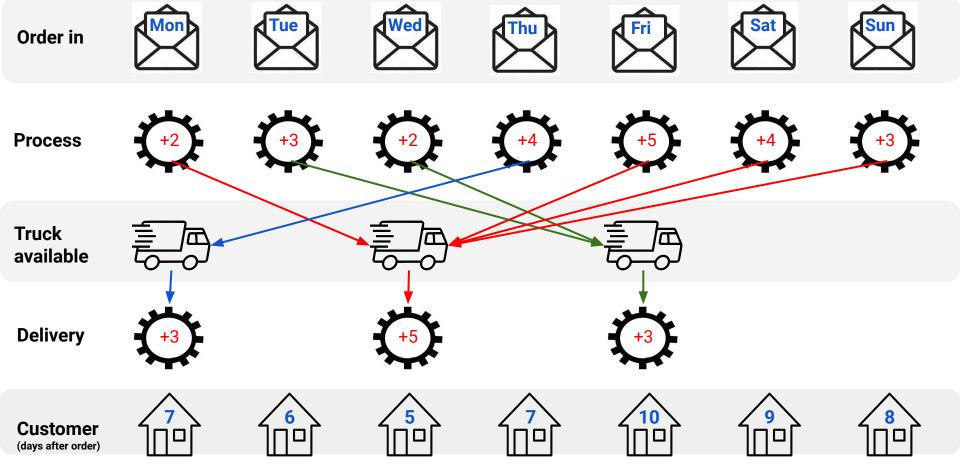


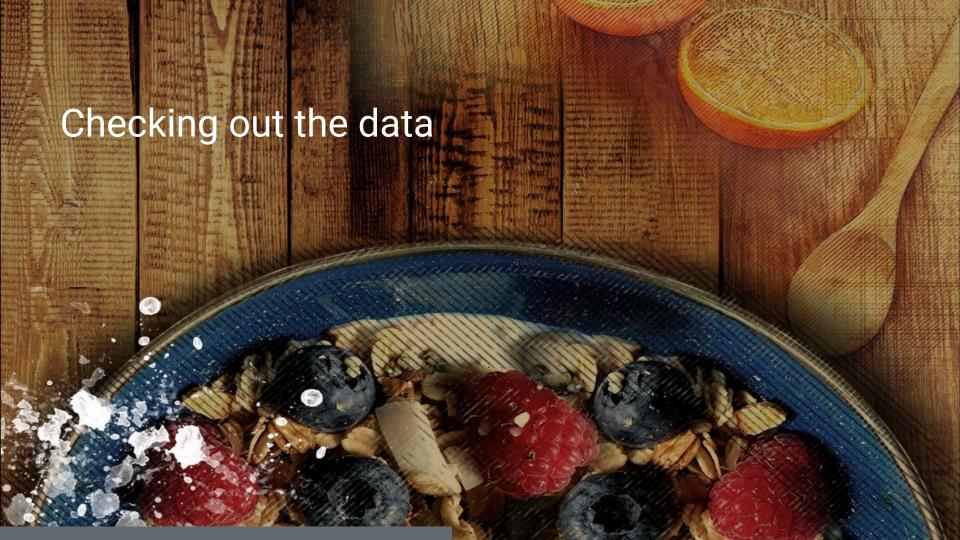




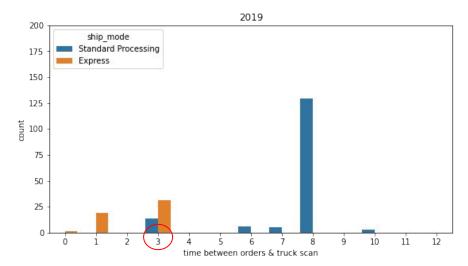


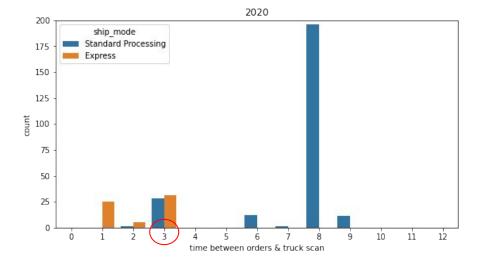






Order on Tuesdays (assumed delay until truck scan = 3_std / 3_exp)



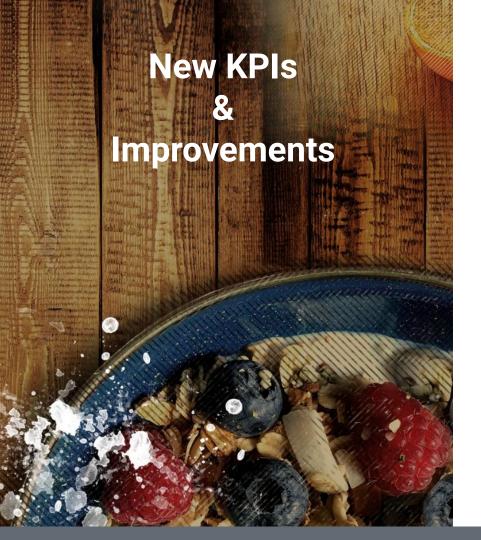




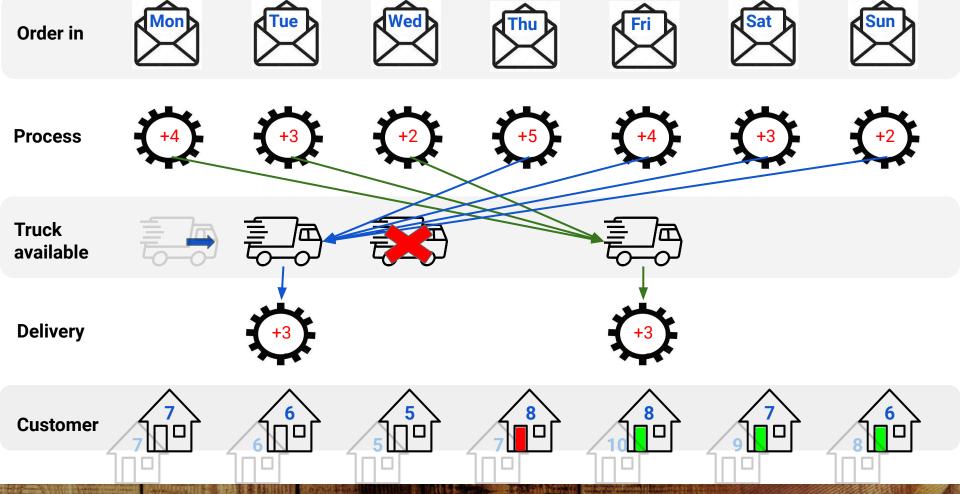




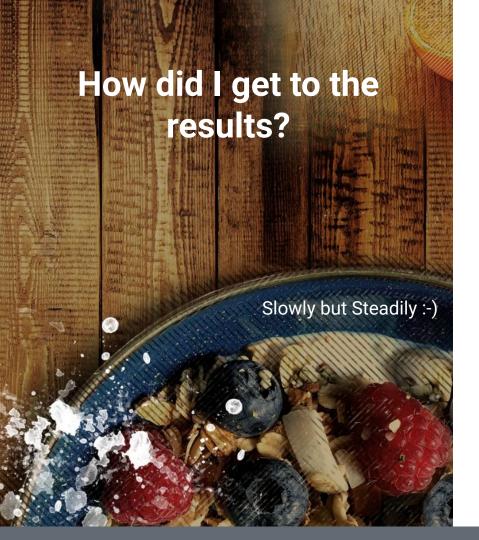
- Real process does not correspond to the assumptions
- Data does not comply with the assumptions
- Process has changed between 2018 and 2019
- Express orders seem to be processed differently than assumed



- Create a process which is tracked constantly
- Track orders with a timestamp (order_date with timestamp)
- Track ready_for_truck with timestamp
- Logistic company track on_truck with timestamp
 - Fixed time slot for package pickup by logistics company
- Describe the process step by step
- Stock overview per product ID







- Creating DataFrame from 'muesli_raw_orders'
 - a. Explore Data
 - b. Clean Data incl. Delete Duplicates
- Creating DataFrame from 'muesli_truck_data'
 - a. Explore Data
 - b. Clean Data incl. Delete Duplicates
- Merge both DataFrames on 'order_id'
 - a. Drop a lot of columns
 - b. Drop columns without truck_data
 - Create columns with workdays and delays between order_date and truck_scan_date
- After Merging and cleaning I just had data for 2 years (2019-2020)
 - Visualise

