Amina Kulshmanova, s0591676, Amina.Kulshmanova@student.htw-berlin.de
Eren Bayoğlu, s0592194, eren.bayoglu@student.htw-berlin.de
Oliver Kwabena Aggrey, s0590504, Oliver.Aggrey@student.htw-berlin.de
Peter Gutjahr, s0590498, peter.gutjahr@student.htw-berlin.de
Nadine Dawaghreh, s0590503, Nadine.Dawaghreh@Student.HTW-Berlin.de

OBETA Warehousing Analytics Project

Literature Review and Planning of Research Task

11.11.2023

htw Hochschule für Technik und Wirtschaft Berlin
University of Applied Sciences

1. OBETA Elektro Vertriebs-GmbH

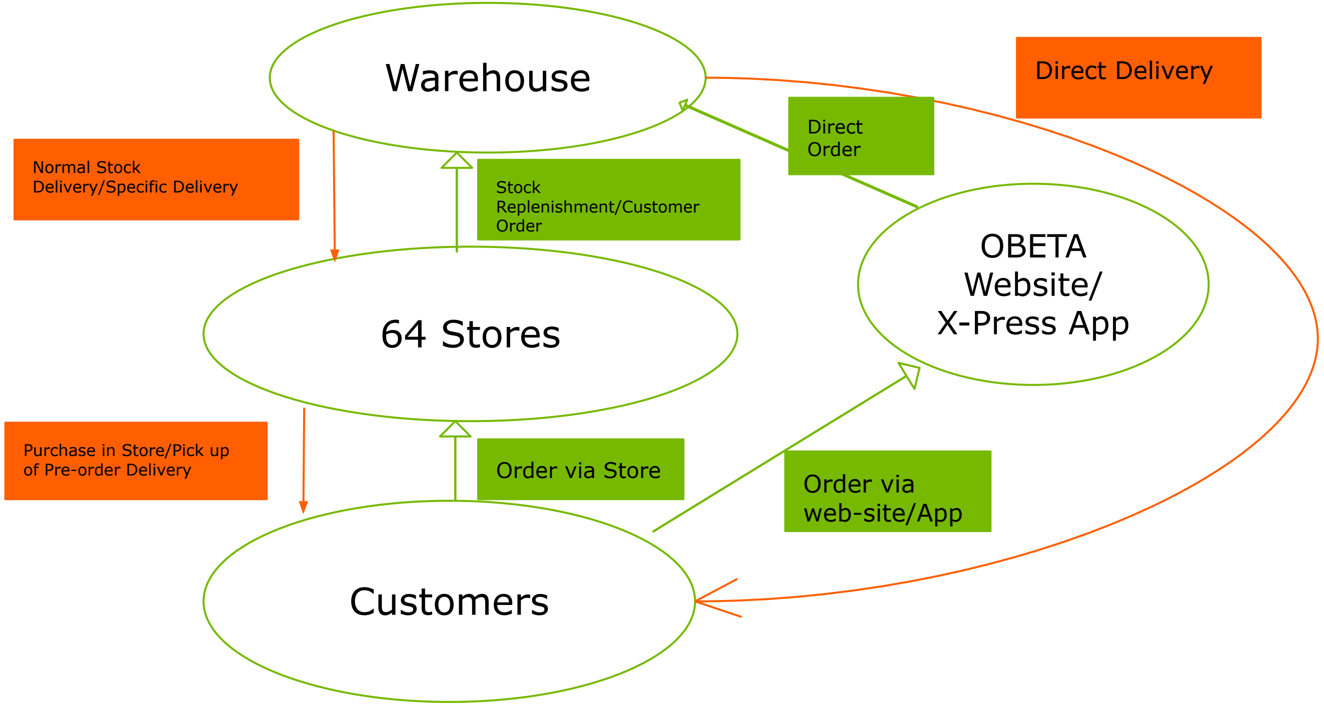
OBETA Business at Glance

- 64 Stores across Germany and X-Press App
- Business Model: B2B, B2C
- Products: Electronics, household appliances, lamps and cables
- Clients: more than 28,000 customers*
- Investments: OSR Shuttle System™, AI Robotic
- Service: Self-service; Delivery and Rental Option

*www. covariant.ai

OBETA’s Order and Delivery Routes

- Order Routes;
- Delivery Routes;



Warehouse at Glance

Warehouse: a storage facility where OBETA's products are **stored** before being **selected** and **dispatched.**

- approx. 48,000 m² including outdoor area
- Five different storage areas
- 24 hour operation (Mon-Fri)
- approx. 20,000 withdrawals/day

Warehouses Processes

Receiving → Putaway → Storing → Picking → Packing → Shipping

Data available in Database

Receiving: Unloading, Identification and Control of delivered Articles

Putaway: Movement of goods from receiving dock to the most optimal storage location in the WH

Storing: Articles are placed in their most appropriate storage space

Picking: Collection of Products to fulfill customer orders

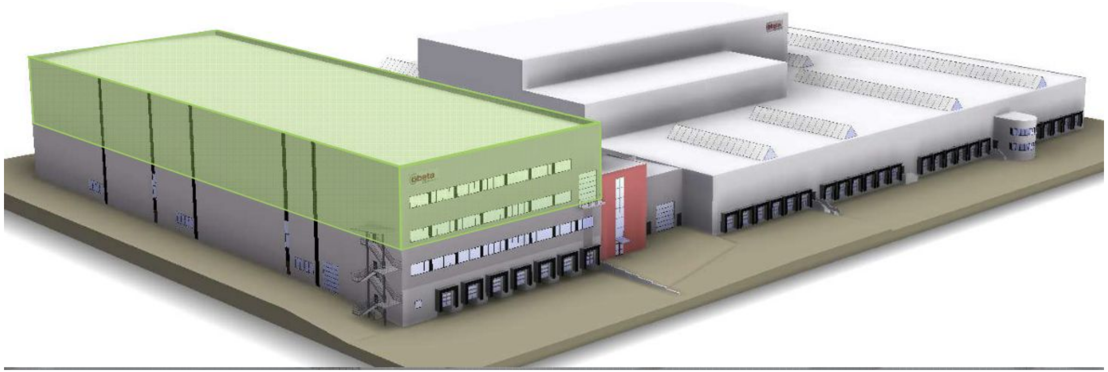
Packing: Consolidating picked items in a sales order and preparing them for shipment

Shipping: Loading and Transport of Products

Data segments within the warehouse

❖ Shuttle Warehouse

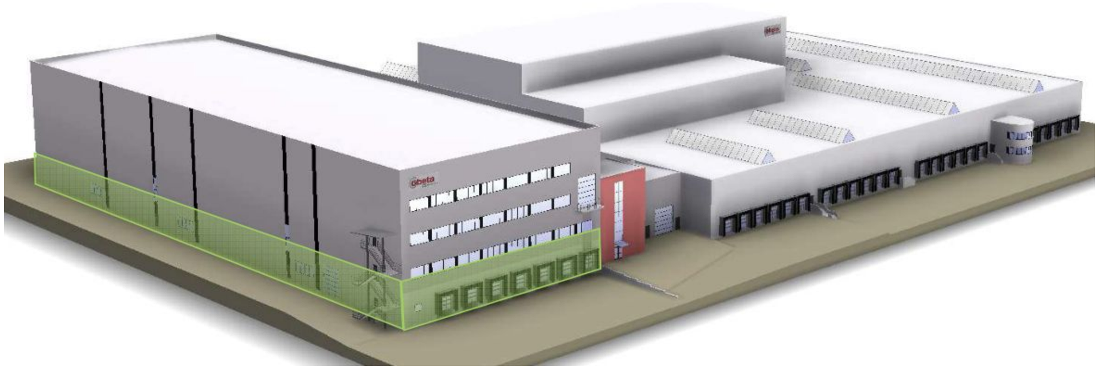
- ➔ Shuttle Robot Systems
- ➔ AI
- ➔ High Performance in terms of picking capacity
- ➔ Extensive product range



Data segments within the warehouse

- ❖ **Cable Warehouse**

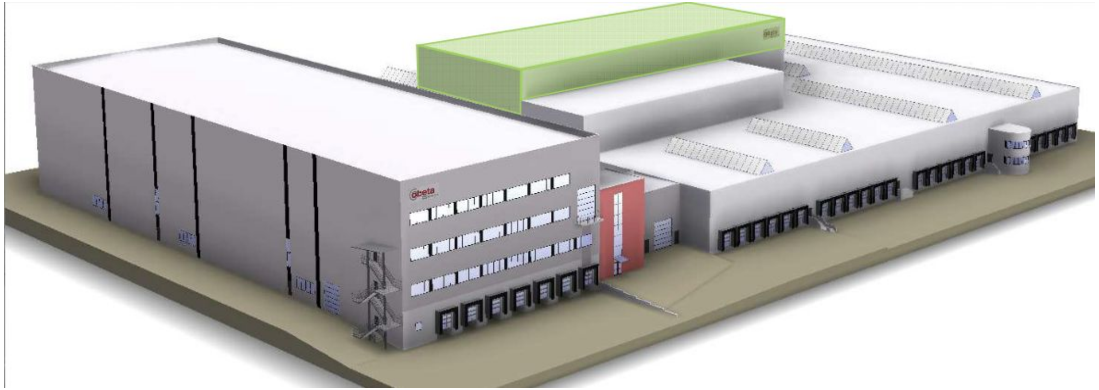
➔ Has lowest activity



Data segments within the warehouse

- **HRL; High-bay Warehouse**

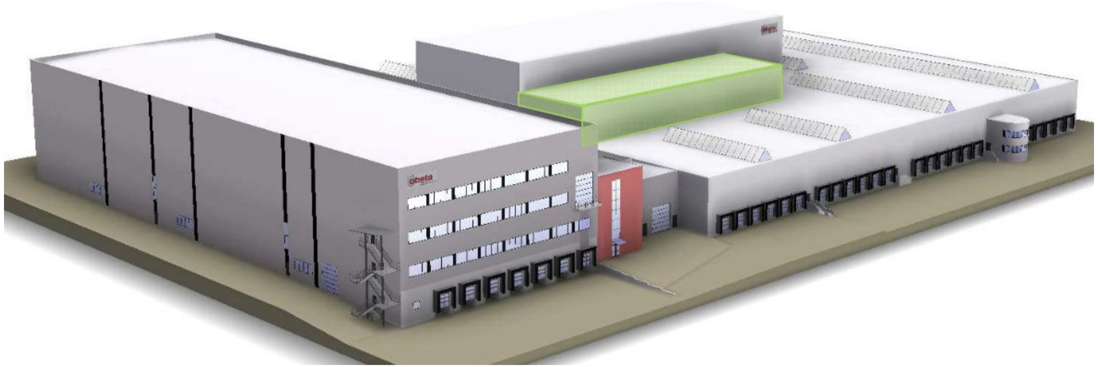
→ High racks
→ High storage density



Data segments within the warehouse

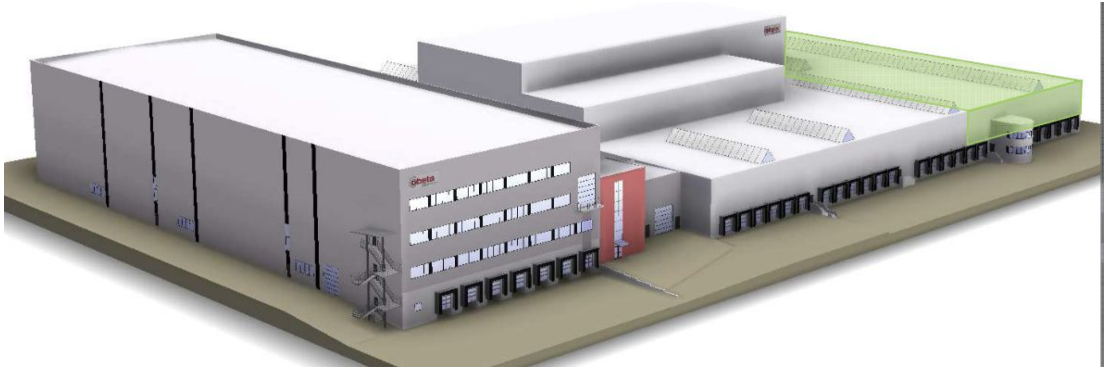
- ❖ **AKL; Automated Small Parts Warehouse**

➔ Small Volume Units



Data segments within the warehouse

- **Manually Operated Warehouse**



2. KPIs

2.4 KPI’s

Category 1 - Warehouse Status

Efficiency of Picks (1. KPI)
A. Total Number of Picks per Timeframe (Day/Week/Month/Year)
B. Avg Time per Pick (per Section)
C. Amount of Unique Products (Years)

Warehouse Section Activity (2. KPI)
A. Number of Picks per Warehouse Section (Month/Year)

Keynote → Status Quo - Inventory, Efficiency

2.4 KPI’s

Category 2 - Order Characteristics

Product Demand (3. KPI)
A. Most wanted vs. Least wanted Products (Year/Month)
B. Picks per Product Group (Year/Month)
C. Stock of unique Products

Customer Ordering Behaviour (4. KPI)
A. Store-Based Orders vs. Direct Warehouse Orders (Year)

Keynote → Status Quo - Satisfaction, Demand, Sales

3. DATA MODELS

Data models

- Data models are utilised for the creation of a simplified and logical database.
- A framework of relationships between the data stored within databases.
- Ensuring consistent definition and formatting of database contents across systems.

Primary data model types

Relational

Dimensional

Entity-Relationship

Entity-Relationship Models (ERM)

Pros
Simplicity of concept
Clear visual representation
Flexibility in modification
Scalability
Relational model integration

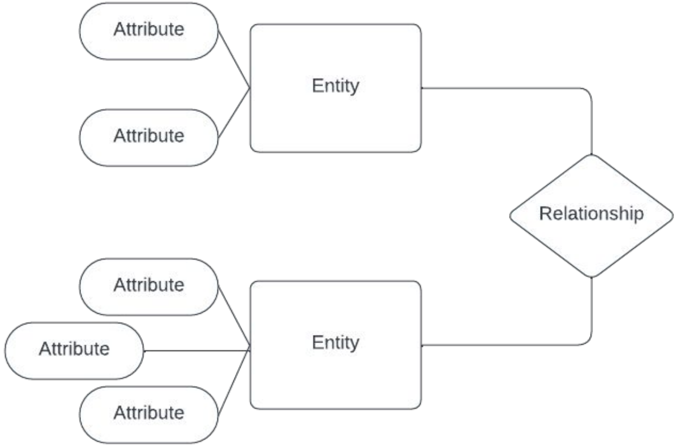


Fig.1: Simple template of ERM

Cons
No industry standard
Possible complexity
Limited to relational models
Optimisation concerns
Assumption of Stability

Star VS Snowflake models

Star Model	Snowflake Model
Dimensions hierarchies are stored in the dimensional table.	Hierarchies are divided into separate tables.
Simple database design.	Very complex database design.
A fact table surrounded by dimension tables.	One fact table surrounded by dimension table which are surrounded by dimension table
Single join only creates relationship between the fact table and any dimension tables.	Many joins required to access data.
Single Dimension table contains aggregated data.	Data Split into different Dimension Tables.

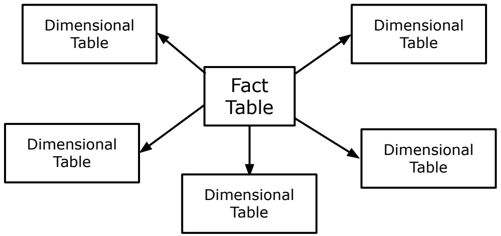


Fig.2: Star Model

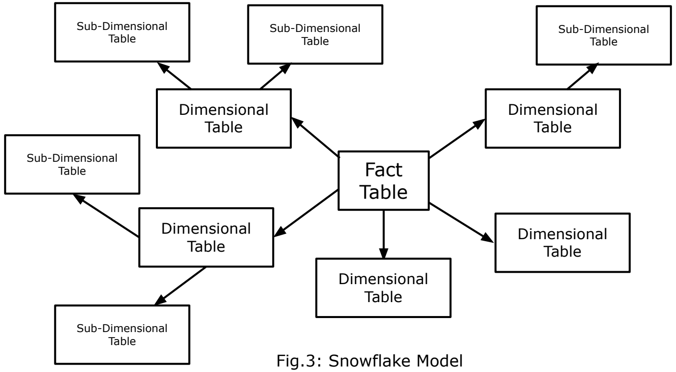
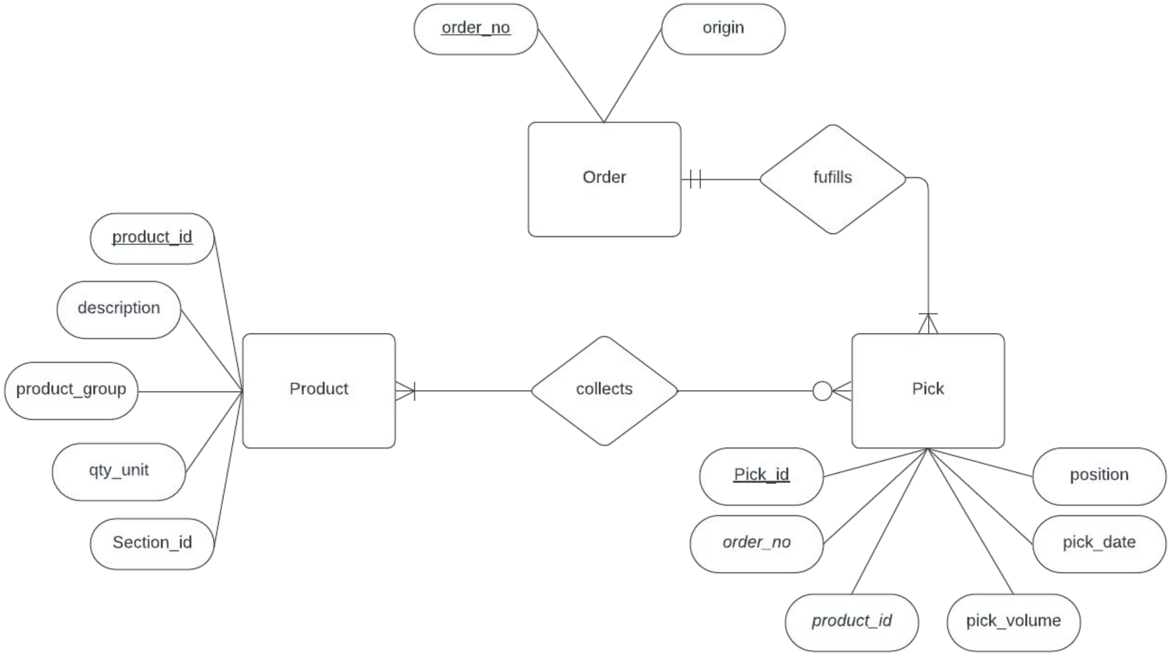


Fig.3: Snowflake Model

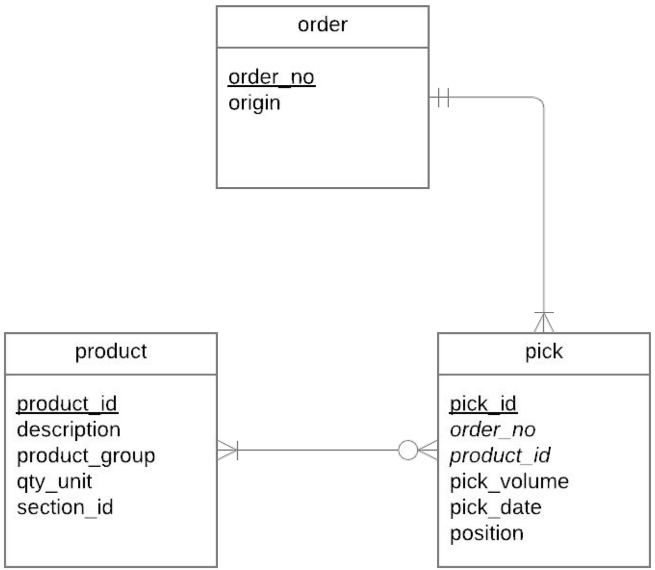
OBETA ERM


order_no
origin
Order
fufills
product_id
description
product_group
qty_unit
Section_id
Product
collects
Pick
Pick_id
order_no
product_id
pick_volume
position
pick_date

OBETA ER Diagram

The three entities are linked through two main relationship:

- Order has **one-to-many** relationship with Pick; mandatory for both.
- Pick has a **many-to-many** relationship with product; product is mandatory while pick is optional.



4. DATA INCONSISTENCIES

Order Numbers: Non-Numeric Values

PROBLEM

Order numbers containing a letter instead of numbers only

SOLUTIONS

Flag the values with "N"

Replace the letter with a number to keep data type consistent

DECISION

These entries likely do not hold any significant value.

Flag these values as "N"

IMPACT

Removal of 25 records has near to zero impact

product_id	warehouse_section	origin	order_number	product_group	description	date
104131	HRL	48	KM00002472	38_Leitungen	NYM-J 5X1 5 100m Mantelltg	2011-09-16 01:15
206452	AKL	48	KM00002249	31_Install.-Befestigungs-Mat.	SPEL AP FR Abzweigds o Kle SD7	2011-07-15 15:13
104131	HRL	48	KM00002242	38_Leitungen	NYM-J 5X1 5 100m Mantelltg	2011-07-15 10:38
104131	HRL	48	KM00002230	38_Leitungen	NYM-J 5X1 5 100m Mantelltg	2011-07-14 14:20
104131	HRL	48	KM00002231	38_Leitungen	NYM-J 5X1 5 100m Mantelltg	2011-07-14 14:20
104131	HRL	48	KM00002232	38_Leitungen	NYM-J 5X1 5 100m Mantelltg	2011-07-14 14:21
104131	HRL	48	KM00002233	38_Leitungen	NYM-J 5X1 5 100m Mantelltg	2011-07-14 14:22
104131	HRL	48	KM00002229	38_Leitungen	NYM-J 5X1 5 100m Mantelltg	2011-07-14 14:22
109910	HRL	48	KM00002226	38_Leitungen	NYM-J 3X1 5 Neu 100m Mantellt	2011-07-14 10:05
104131	HRL	48	KM00002215	38_Leitungen	NYM-J 5X1 5 100m Mantelltg	2011-07-13 20:57
104131	HRL	48	KM00002212	38_Leitungen	NYM-J 5X1 5 100m Mantelltg	2011-07-13 20:57

Fig.1: Data sample of the non-numeric order numbers

- Majority of records were from the same warehouse section, product group, and date frame.
- Since last entry was in 2011, not up to date we can discard them.

Order Numbers: Repeating Values

PROBLEM

There are order numbers repeating across different years.

SOLUTIONS

Make order numbers unique for better track

DECISION

Create a new column named "adjusted_order_number" to make each order unique

IMPACT

For better KPI track, order numbers must be unique for each pick

```
There are order numbers repeating across different years.
        order_number  year  counts
            01000002  2015       1
            01000002  2019       5
            01000002  2020       1
            01000003  2019       6
            01000003  2020       2
                 ...   ...     ...
            96880711  2012      18
            96880881  2011       4
            96880881  2012       3
            98327271  2011       6
            98327271  2012       2

[4660743 rows x 3 columns]
```

Fig.2: Order numbers that are repeating across different years

2.005.663
repeating value records

6%
of total pick records

Missing Values in Merged Data

PROBLEM

Product IDs that have been picked but are not present in the product data

SOLUTIONS

Replace the null value with 'Unknown'

Keep as is, unedited

DECISION

Replace the null value with 'Unknown'

Update the missing values in the product data for future entries

IMPACT

Maintain data integrity and consistency.

Fig.3: Missing Values in Merged Data

product id	warehouse section	origin	order number	position in_order	pick volume	quantity unit	date	description	product group
Y91358	SHL	48	1E+06	2	1	St	2019-07-29 22:12:54		
Y91360	SHL	48	9E+06	3	1	St	2019-03-04 10:38:25		
Y91380	SHL	48	9E+06	2	1	St	2019-03-04 10:27:54		

3
missing value records

Missing Values in Product Data

PROBLEM	SOLUTIONS	DECISION	IMPACT
A number of products missing a product description value	Replace the null value with 'Null' Keep as is, unedited	Keep as is, unedited Later on, make sure product_data is up to date	Description doesn't have a notable impact on desired data analysis because they're not in our pick data

Fig.4: Data sample of missing description values

product_id	description	product_group
1849999		42_Netzwerktechnik
256092		31_Install.-Befestigungs-Mat.
273647		33_Schaltgeräte
276118		33_Schaltgeräte
276121		33_Schaltgeräte
276344		33_Schaltgeräte
296471		33_Schaltgeräte
330269		35_Leuchten
363612		37_Leuchtmittel
380696		40_Sprechen_Schwachstrom
382025		41_Antenne
388498		41_Antenne

6.023
missing value records

0,3%
of total product records

Duplicated Values

PROBLEM

Duplicate rows

SOLUTIONS

Flag the values with “N”

DECISION

Remove the flagged records of duplicates

IMPACT

Removal of 8.024 records has near to zero impact

Fig.5: Data sample of missing description values

product id	warehouse section	origin	order number	position in_order	pick volume	quantity unit	date	description	product group
2	SHL	48	7055448	1	22	St	2017-06-30 12:50:40	Obo T 60 M20 NL Kabelabzweig	31_Install.-Befestigungs-Mat.
2	SHL	48	7055448	1	22	St	2017-06-30 12:50:40	Obo T 60 M20 NL Kabelabzweig	31_Install.-Befestigungs-Mat.
12	AKL	48	8197426	3	4	St	2018-01-30 12:55:47	ETAP D42R1/LEDN1820SV130	35_Leuchten
12	AKL	48	8197426	3	4	St	2018-01-30 12:55:47	ETAP D42R1/LEDN1820SV130	35_Leuchten
99	SHL	48	3302027	1	4	St	2020-07-14 07:00:04	Spels 12740816 TKPC2518-9-m	34_Verteiler_Schränke
99	SHL	48	3302027	1	4	St	2020-07-14 07:00:04	Spels 12740816 TKPC2518-9-m	34_Verteiler_Schränke
255	Manuell	48	8883282	1	8	St	2018-11-21 00:33:42	.BX1OSE-830M-D970	35_Leuchten
255	Manuell	48	8883282	1	8	St	2018-11-21 00:33:42	.BX1OSE-830M-D970	35_Leuchten

8.024
duplicated value records

0,02%
of total pick records

Pick Volume: Negative Values

PROBLEM

Records display a negative values for "**pick_volume**"

SOLUTIONS

Flag the values with "N"

Transform value to absolute value assuming human error

DECISION

Remove the flagged records with negative values.

IMPACT

Removal of 100 records has near to zero impact

product_id	warehouse_section	origin	order_number	position_in_order	pick_volume	quantity_unit	date	description	product_group
R12006	SHL	48	1670127	4	-3	St	2015-09-04 11:32	HAGER Schütz 40A 4S bf ESC440S	33_Schaltgeräte
294723	SHL	48	1672819	6	-1	St	2015-09-04 13:06	SPEL Kleinvert 12TE IP65 AK12	34_Verteiler_Schränke
234142	SHL	48	1723719	9	-25	St	2015-09-15 16:32	WAGO COMP Verb-Klemme 5x0 14-4	31_Install.-Befestigungs-Mat.
206454	SHL	48	1728544	1	-10	St	2015-09-16 14:35	SPEL AP FR Abzweigds o Kle I12	31_Install.-Befestigungs-Mat.
R20556	SHL	48	1737252	6	-1	St	2015-09-17 13:14	LEGR NILOE 2f Rahmen uw 665002	32_Schalter_Steckvorrichtg
250212	SHL	48	1800572	1	-10	St	2015-09-30 10:34	B/J REFL Steckdos aw 20EUC-214	32_Schalter_Steckvorrichtg
310053	SHL	46	1806800	5	-1	St	2015-10-01 07:54	RZB Nurglasl 60W E27 10120.002	35_Leuchten
K85000	SHL	46	1806800	2	-10	St	2015-10-01 07:54	LEGR PLEXO Steckd 1f gr 069730	32_Schalter_Steckvorrichtg

Fig.6: Data sample of negative pick volume

Pick Volume: 0 Values

PROBLEM

Records has a "0" values for **"pick_volume"** representing 0 products picked for that instance

SOLUTIONS

Flag the values with "N"

Identify which orders contained 0 pick volumes and identify possible scenarios

DECISION

Remove the flagged records of zero pick values

Use flagged data to analyse for future

IMPACT

Removal of 190.272 records has minimal impact

product_id	warehouse_section	origin	order_number	position_in_order	pick_volume	quantity_unit	date
32	AKL	48	6183218	18	0	St	2017-01-17 00:32
32	AKL	48	6183218	18	0	St	2017-01-17 00:35
77	AKL	48	7746735	3	0	St	2017-11-08 22:43
104	AKL	48	8221188	2	0	St	2018-02-01 21:34
104	AKL	48	8221188	2	0	St	2018-02-01 21:44
224	AKL	48	8447450	1	0	St	2018-09-14 05:42
100500	HRL	48	9603326	6	0	Mt	2019-03-20 02:19
100503	HRL	48	1422908	21	0	Mt	2019-08-06 18:09
100699	AKL	48	8122443	1	0	Mt	2018-01-17 08:56
100700	AKL	48	1765638	1	0	Mt	2019-10-08 14:41

Fig.7: Data sample of 0 pick volume

190.272
0 value records

0,56%
of total pick records

5. Outliers

Handling Outliers

- Outliers are data points that significantly differ (**are extreme**) from majority of the data in a dataset.

obs:	pick_v~e	wareho~n	order_no	date	origin
1.	-2000	SHL	02542493	2016-02-22 15:02:34.0000000	46
2.	-200	SHL	03017429	2016-05-25 02:46:32.0000000	48
3.	-150	SHL	03268633	2016-07-11 20:34:37.0000000	48
4.	-150	SHL	03239830	2016-07-06 01:19:47.0000000	48
5.	-100	SHL	03035238	2016-05-26 22:49:57.0000000	48

3.39e+07.	100000	HRL	06312388	2013-07-16 11:25:45.0000000	48
3.39e+07.	100000	HRL	07981876	2017-12-14 16:04:33.0000000	48
3.39e+07.	107000	HRL	05148780	2018-03-26 14:45:40.0000000	48
3.39e+07.	152000	HRL	05535879	2018-06-07 18:07:10.0000000	48
3.39e+07.	200000	Manuell	03417870	2012-05-24 17:26:52.0000000	48

Fig.10: Extreme data points

- Interquartile range (IQR) to detect outliers based on normality criteria.
- Used data visualization techniques like box plots to visually identify outliers in data.

Overall 11,94%
outliers within product groups

Impact
Are outliers legitimate data points or the result of errors or rare events

Handling Outliers

Classification of Data Inconsistencies as Outliers

- Addressing these inconsistencies as a crucial initial step in the analysis process.

Anderson-Darling Test for Normality on 'Pick Volume' Variable

- Conducting this test to determine the appropriate outlier detection method.
- Strong rejection of the assumption of normal distribution.
- Decision to use the Interquartile Range (IQR) method for outlier detection based on this result.

Preparation and Implementation of the IQR Method

- Implementing the IQR method as part of the data analysis process.

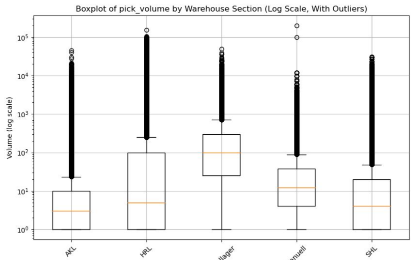


Fig.8: Boxplot of pick volume by Warehouse Section

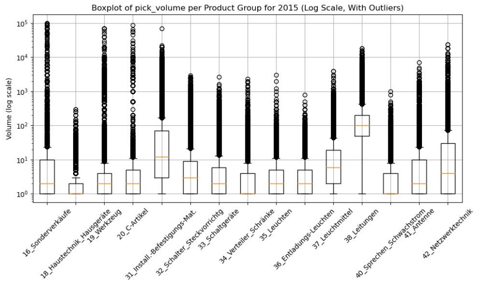


Fig.9: Boxplot of pick volume per Product Group in 2015

Impact of Properly Dealing with Outliers:

- Importance of addressing outliers: they can significantly skew analysis results and lead to incorrect conclusions.
- Correct handling of outliers ensures accurate representation of central trends in the dataset.
- Identifying outliers can reveal potential errors in data collection or provide insights into unusual but significant occurrences within the dataset.

Summary of Findings and Next Steps

Operational Summary				
AKL ***(Automated Small Parts)***	**SHL** ***(Automated Shuttle)***	**HRL** ***(Bulk Items)***	**Manual Warehouse**	**Kabelager** ***(Cables)***
Extensive unique orders with high automation	Peak in orders and volume, highly automated	Lower order frequency, large-volume handling	High manual handling, varied order sizes.	Moderate orders, specialized handling needs

Efficiency Takeaways

- Automation drives AKL and SHL efficiency
- HRL and Manual Warehouse adapt to item size and order variability
- Kabelager relies on custom handling

Next Step

- Advanced Data Analysis and Refinement
- Enhanced Visualization and Dashboard Development
- Actionable Insights and Strategic Recommendations

Hochschule für Technik
und Wirtschaft Berlin

University of Applied Sciences

www.htw-berlin.de

Eren Bayoğlu. s0592194. eren.bayoglu@student.htw-berlin.de
Oliver Kwabena Aggrey. s0590504. Oliver.Aggrey@student.htw-berlin.de
Peter Gutjahr. s0590498. peter.gutjahr@student.htw-berlin.de
Nadine Dawaghreh. s0590503. Nadine.Dawaghreh@Student.HTW-Berlin.de

OBETA Warehousing Analytics Project

Data Cleansing and Univariate Data Analysis

15.12.2023

Hochschule für Technik und Wirtschaft Berlin
University of Applied Sciences

Index

Current State

- First Presentation
  Business Understanding
- Data Understanding
- Selection of KPIs
- ER Model
- Inconsistencies, Missing Values, Outliers
- Second Presentation
  Deep Data Understanding
- Univariate Analysis
- Assessment of KPI`s
- Data Preparation
  Data Cleaning & Transformation (ETL)
- Documentation of ETL Process
- Business Understanding
  Knowledge about OBETAS Performance
- Next Steps

1. ETL Process

Process Summary
2% of the data have been removed
30% of the data have been edited/replaced

1.1 Data Cleaning Process

FINDING/PROBLEM	OPTIONS TO SOLVE THE ISSUE	DECISION MADE	IMPACT OF DECISION	RECORDS COUNT
Order numbers containing a letters	• Flag values as "N" • Replace letter with number	Flag rows and remove them	25 records are affected; little impact	Before: 33.888.990 After: 33.888.965
Repetitive order numbers in 4.6M rows	• Adjust order numbers to prevent repetition	Addition of column with new order numbers	Improved KPI tracking with unique orders; 2.005.663 new unique orders numbers	Before: 33.888.965 After: 33.888.965
Products description and product group missing in product data	• Replace the null value with 'Unknown' • Keep as is. unedited	Replace the null value with 'Unknown'; Update missing values in the product data	Only 3 records. however. Maintain data integrity and consistency.	Before: 33.888.965 After: 33.888.965
0 values for Pick Volume	• Flag the values with "N" • Identify possible scenarios for 0 pick volume	Remove the flagged records of zero pick values;	Removal of flagged records; removal of 190.272 records has minimal impact	Before: 33.888.965 After: 33.698.694
Negative values for Pick Volume	• Flag the values with "N" • Transform to absolute value; human error	Remove flagged records with negative values	With only 100 records. removal has little impact on the overall analysis	Before: 33.698.694 After: 33.698.594
A number of duplicated rows	• Flag rows as duplicates • Retain the duplicates in database	Flag rows as duplicates and remove them	Removal of 8.024 duplicate pick data provides a more accurate analysis	Before: 33.698.594 After: 33.690.589
Missing product description	• Replace the null value with 'Null' • Keep as is. unedited	Keep as is. unedited	Description doesn't have a notable impact on desired data analysis because they're not in our pick data	Before: 33.698.594 After: 33.690.589
Special German Characters	• Replacing with English characters	Replacing with English characters	Number of rows changed in 'description': 3584575 Number of rows changed in 'product_group': 5581557	Before: 33.698.594 After: 33.690.589
Outlier Detection	•Iterate over each group to calculate the 3-sigma interval and identify outliers	Flag the values with "N" and remove	A total of 409.124 rows deleted	Before: 33.690.589 After: 33.281.465
Products with a non-unique quantity unit	• Use unique quantity unit • Flag the values with "N"	Flag the values with "N" and remove	A total of 3.918 rows belonging to 71 products deleted. The effect on the total population is negligibly small.	Before: 33.281.465 **After: 33.277.547**

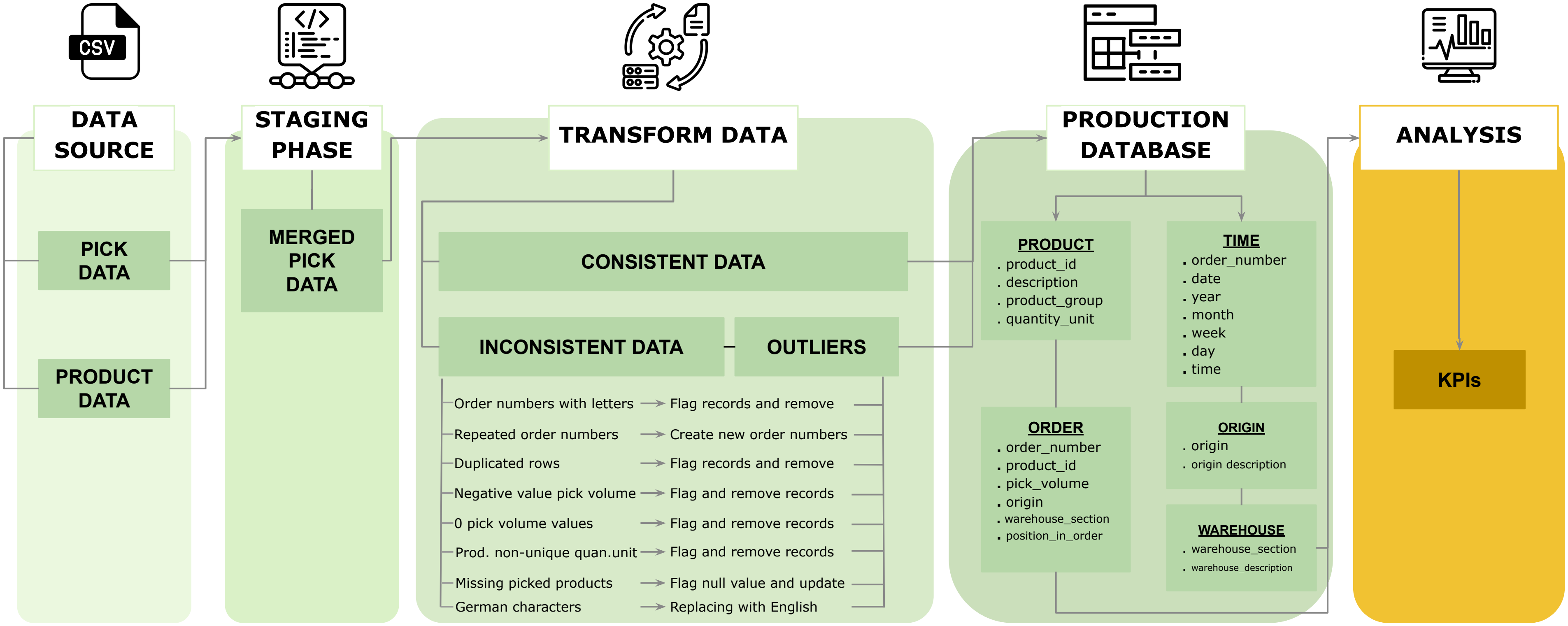
1.2 Effective Outlier Detection: OBETA's Data Strategy

Statistical Outlier Analysis Based on Product Group and Year

- Analyzing outliers by product group and year is vital for OBETA, as it allows for the detection of year-specific trends and operational anomalies within each unique product category.
- This targeted analysis is essential for effective inventory management, enabling OBETA to adapt to market conditions and optimize processes on a per-product-group basis annually.

Year	2019				
Row Labels	**Average of Mean Volume**	**Average of Standard Deviation**	**Average of Lower Bound**	**Average of Upper Bound**	**Sum of Number of Outliers**
31_Install.-Befestigungs-Mat.	99.0	317.7	-854.2	1052.2	13562
32_Schalter_Steckvorrichtg	12.0	41.4	-112.3	136.3	10214
33_Schaltgeraete	8.4	25.3	-67.6	84.3	6309
38_Leitungen	223.8	530.7	-1368.3	1816.0	4686
41_Antenne	41.3	141.2	-382.4	465.1	3447
37_Leuchtmittel	15.5	31.8	-79.9	110.9	3118
19_Werkzeug	113.1	620.2	-1747.6	1973.8	1973
40_Sprechen_Schwachstrom	4.9	13.1	-34.4	44.2	1525
42_Netzwerktechnik	100.7	366.5	-998.7	1200.1	1445
35_Leuchten	6.4	25.5	-70.1	82.8	1121
36_Entladungs-Leuchten	4.8	8.8	-21.5	31.1	566
34_Verteiler_Schraenke	6.5	75.2	-219.0	232.1	323
18_Haustechnik_Hausgeraete	2.5	8.1	-21.9	26.9	297
20_C-Artikel	185.3	2093.1	-6094.1	6464.7	121
16_Sonderverkaeufe	247.6	3070.9	-8965.2	9460.5	58
Unknown	1.0	0.0	1.0	1.0	0
Grand Total	**67.1**	**460.6**	**-1314.8**	**1448.9**	**48765.0**
Total Outliers of All Years	**74**	**467**	**-1,327**	**1,475**	**409,124**

1.3 ETL Process



1.4 Univariate Analysis

Count for all variables: 33.277.547

Variable Name	Variable Type	Unique	Top	Frequency
product_id	Nominal	96.822	‘109910’	171.441
description	Nominal	88.872	NYM-J 3X1 5 Neu 100m Mantelltg	171.504
quantity_unit	Nominal	9	St	28.118.872
warehouse_section	Nominal	5	SHL	14.385.452
product_group	Nominal	16	NYM-J 3X1 5 Neu 100m Mantelltg	8.750.223
order_number	Nominal	9.319.844	‘40557142016’	371
origin	Dichotomous	2	48	26.144.017

Variable Name	Variable Type	Mean	Min	%25. 50. 75	Max	Std
origin	Dichotomous	-	46	48	48	0.82
position_in_order	Ordinal	5.42	1	1 - 3 - 7	436	6.94
pick_volume	Metric	46	1	1 - 5 - 20	11.000	141.89
date	Interval	2016-07-01 03:33:12	2011-06-23 00:00:01	2014-05-07 06:04:14. 2016-10-24 18:40:49. 2018-09-28 12:56:47	2020-07-14 11:42:01	-

2. KPIs

2.1 Breakdown on different Levels

	Warehouse Status (Picks)	Order Characteristics (Over Time)
Primary KPIs	1. Number of Picks over Time	7. Number of Incoming Orders (over Time)
	2. Number of Picks per Warehouse Section	8. Number of Orders per Origin (46.48)
	3. Number of Picks per Product Group	9. Seasonal Order Variation (Years)
	4. Number of Picks per Product (Top)	10. Duration to complete an Order
	5. Number of unique Products	11. Top products ordered
Secondary KPIs	6. Average Picks per Day/Week/Month	12. Average Products per Order
		13. Time to complete an Order (per Section)

2.2 KPI Hierarchy

3. Number of Picks per Product Group
4. Number of Picks per Product (Top)
9. Seasonal Order Variation (Years)
11. Top products ordered

Top Tier (Strategic KPIs)

1. Number of Picks over Time
6. Average Picks per Day/Week/Month
7. Number of Incoming Orders (over Time)
8. Number of Orders per Origin (46. 48)
12. Average Products per Order

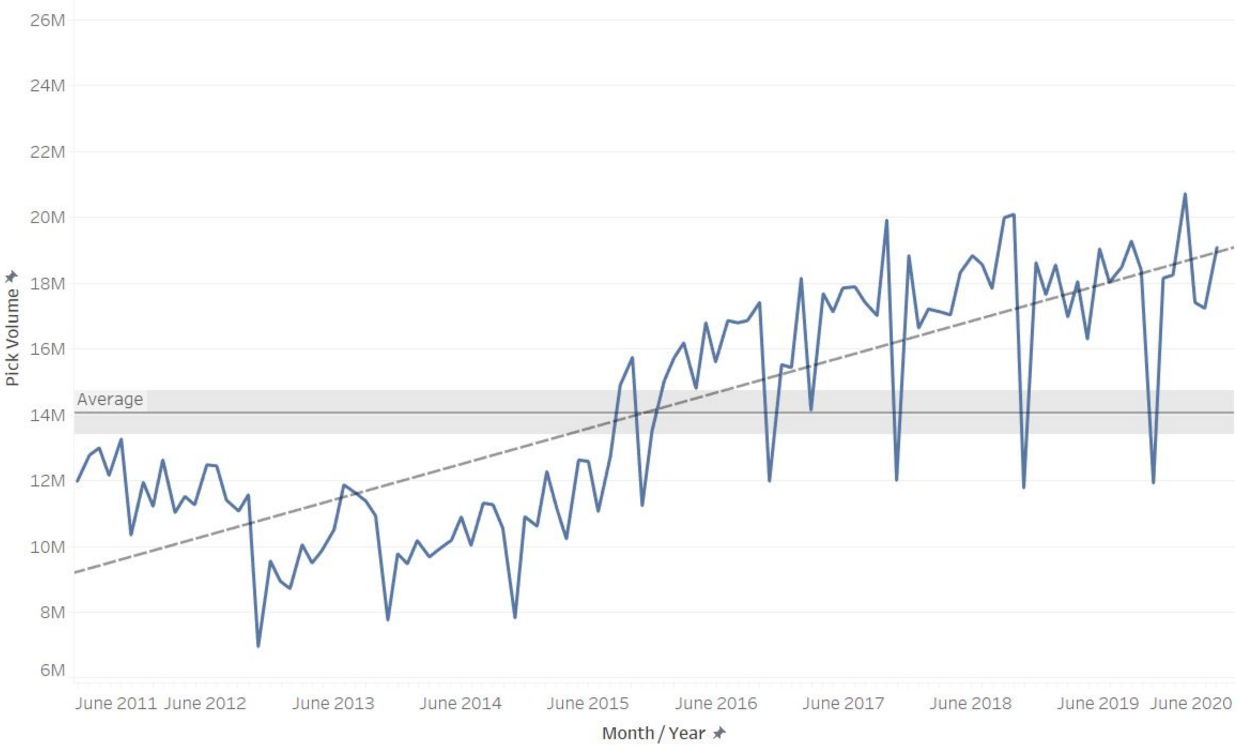
Mid-Level (Tactical KPIs)

2. Number of Picks per Warehouse Section
5. Number of unique Products
10. Duration to complete an Order
13. Time to complete an Order (per Section)

Base Level (Operational KPIs)

2.1 Warehouse Status (Picks)

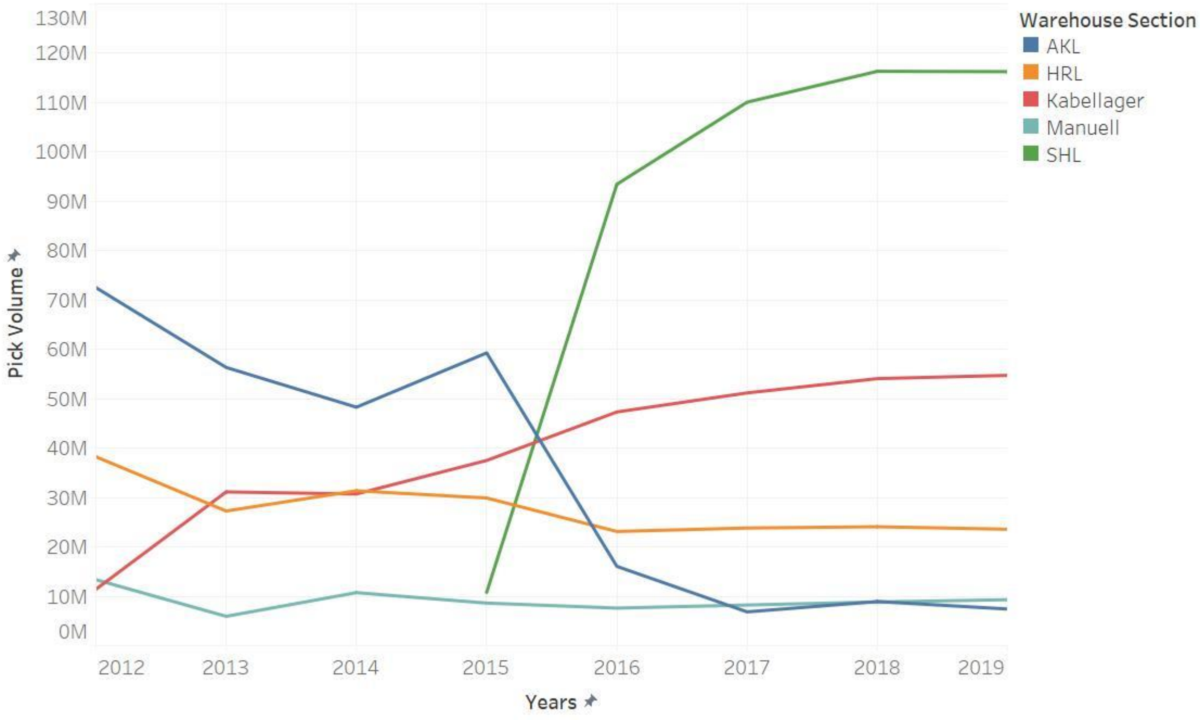
2.1.1 Overall Pick Volume



Insights

- There had been an increase in pick volume starting in **2015** impacted by launch of **SHL**.
- **December** consistently recorded the lowest pick volume for the year, scoring lower than the average of the reported period. to recover in January.
- The end of year and holidays influences the pick volume during these periods.

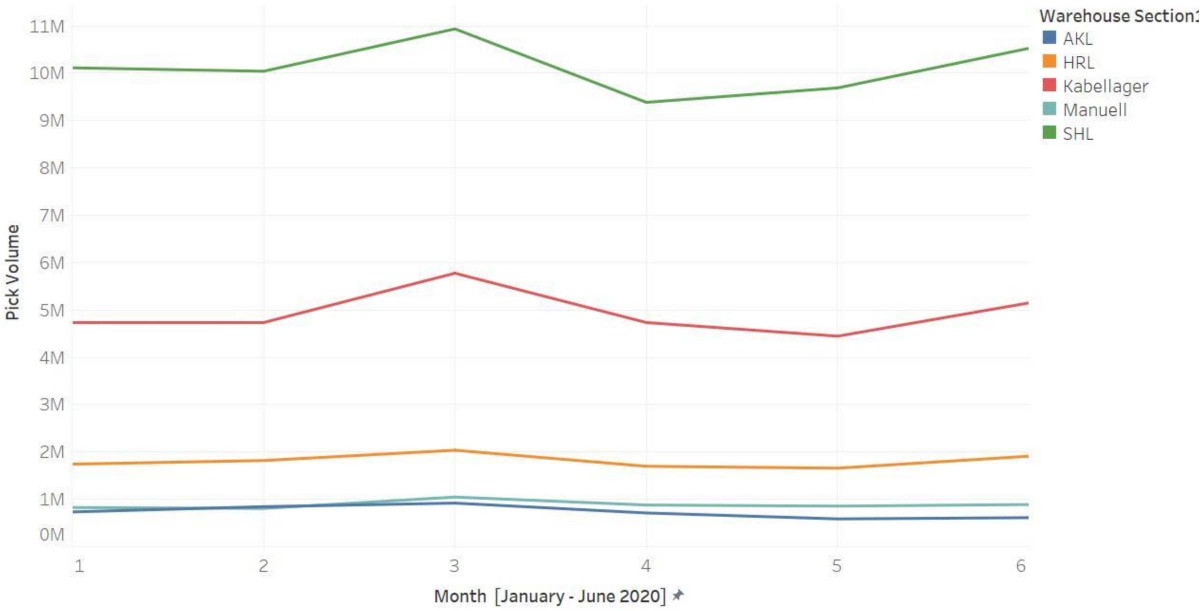
2.1.2 Number of Picks per Warehouse Section



Insights

- **SHL** was integrated in the year 2015.
- **AKL** recorded the highest pick volume until the introduction of **SHL**.
- **SHL** overtook **AKL** within the year and now holds the highest pick volume.
- Incomplete data in 2011 & 2020

2.1.3 Number of Picks per Warehouse Section (2020)

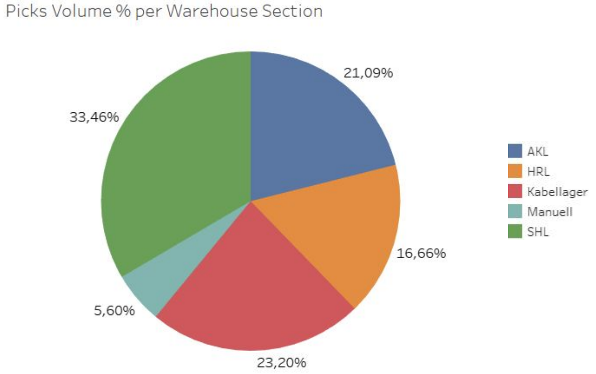


Insights

- **AKL** and **Manuell** are equalized in terms of pick volume.
- **SHL** is constantly retaining the highest pick volume.

2.1.4 Number of Picks per Warehouse Section

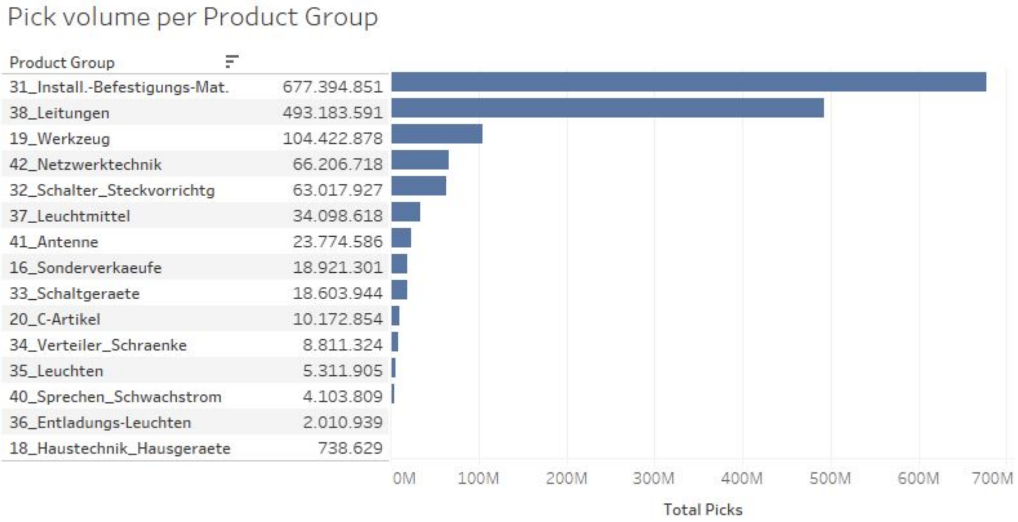
Warehouse Sections Ranking per pick volume	
512M	SHL
355M	Kabellager
323M	AKL
255M	HRL
86M	Manuell



Insights

- SHL saw the **highest pick volume**. despite a later launch in 2015.
- SHL generated **33,5%** of overall pick volume recorded.
- Kabellager followed, in which 21,1% of pick volume was conducted.
- Manuell saw the **least pick volumes**, making 5,6% of overall volume.

2.1.5 Number of Picks per Product Group

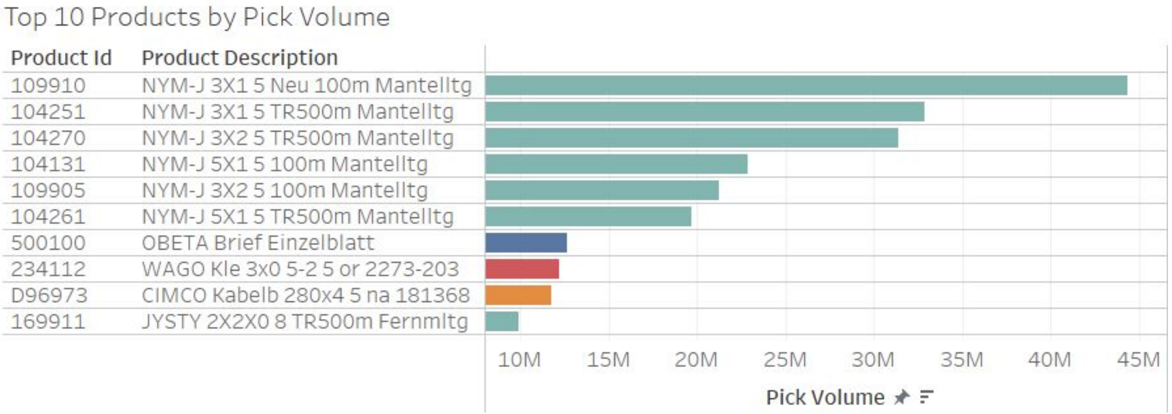


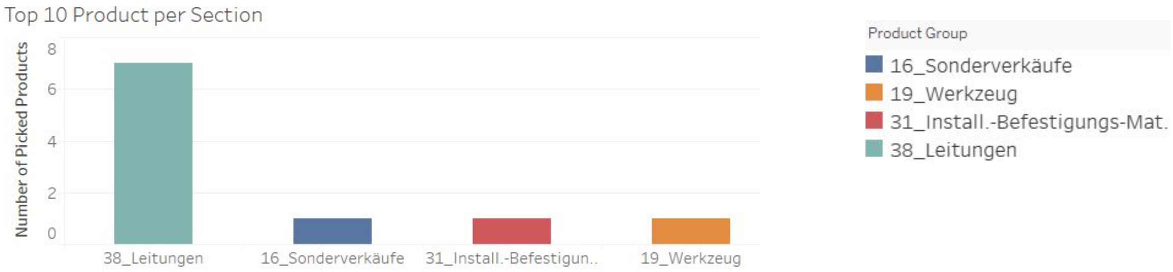
Insights

- **Product Group 31** achieved the highest pick volume (44,3% of total volume)
- **Product Group 18** ranked last in pick volume, making up only 0,05%.

31	38	19	42	32	37	41	16	33	20	34	35	40	36	18
44,25%	32,22%	6,82%	4,33%	4,12%	2,23%	1,55%	1,24%	1,22%	0,66%	0,58%	0,35%	0,27%	0,13%	0,05%

2.1.6 Top 10 Products Picked





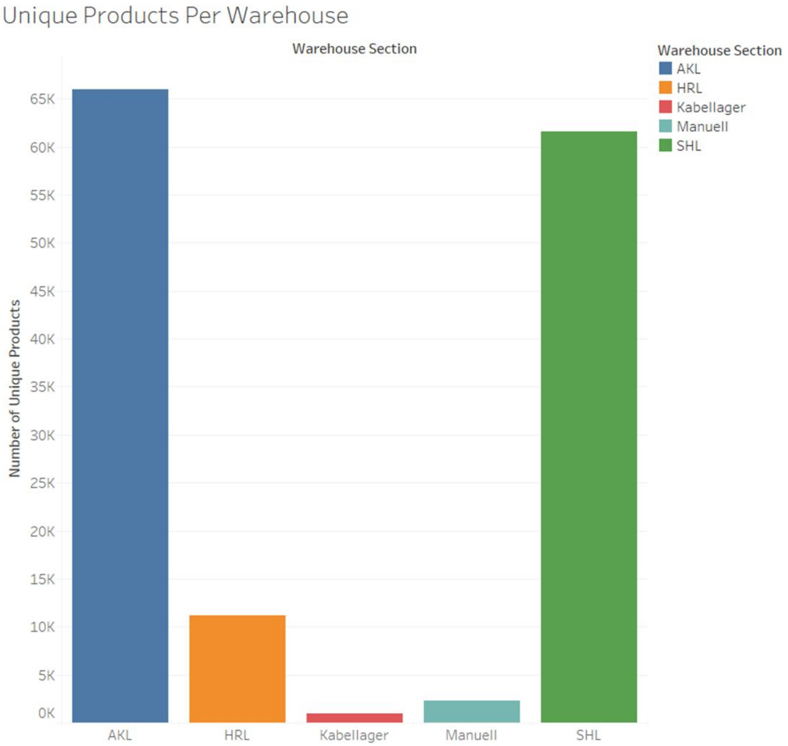
Insights

- Majority of the top 10 picked products are from **product group 38** with 7 products recording over 182,3M picks.

2.1.7 Unique Products

Warehouse Section	
AKL	66,016
SHL	61,624
HRL	11,187
Manuell	2,299
Kabellager	956

Sum of Number of Unique Products broken down by Warehouse Section.

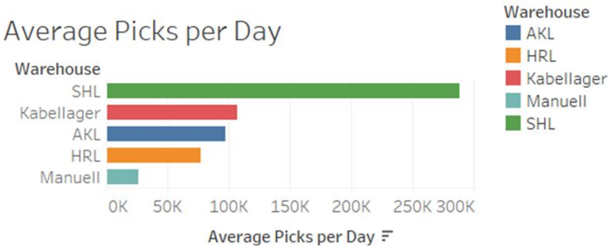


Insights

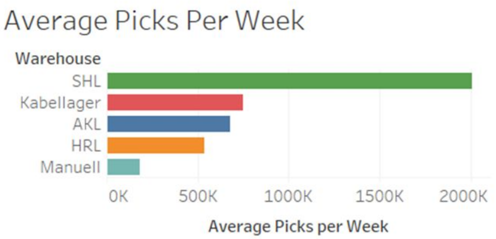
- **AKL** warehouse has the highest number of unique products with **66.016** products.
- **AKL** and **SHL** combined store **89.9%** of total unique products.
- **Cable** warehouse has the lowest number of unique products with **956** products.

2.1.7 Average Picks per Day/Weeks/Month

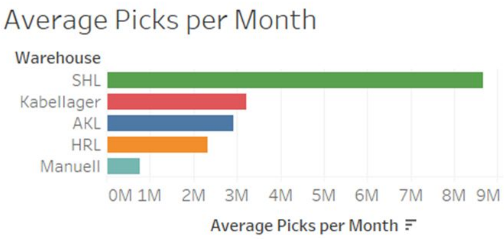
Warehouse	
SHL	287,882
Kabellager	107,274
AKL	97,527
HRL	77,055
Manuell	25,887



Warehouse	
SHL	2,008,403
Manuell	180,774
Kabellager	749,106
HRL	538,086
AKL	681,045



Warehouse	
SHL	8,680,388
Manuell	778,973
Kabellager	3,227,966
HRL	2,318,660
AKL	2,934,683



Insights

- **SHL** warehouse has the highest average picks across time, with an average of:

287.882	picks/day
2.008.403	picks/week
8.680.388	picks/month

2.2 Order Characteristics (Over Time)

2.2.1 KPI Number of Incoming Orders

Volume

6.928.668

Number of unique Orders
BEFORE
Differentiating Duplicates

9.319.844

Number of unique Orders
AFTER
Differentiating Duplicates

Efficiency

3.479
Avg. Orders per
Day
excluding non working days

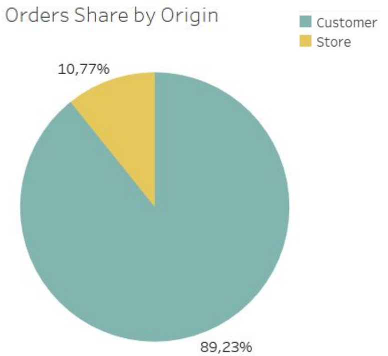
19.778
Avg. Orders per
Week

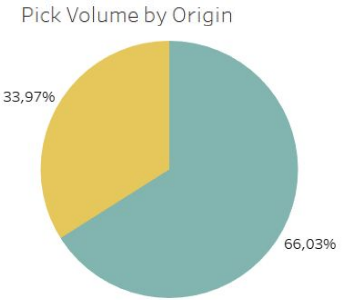
84.865
Avg. Orders per
Month

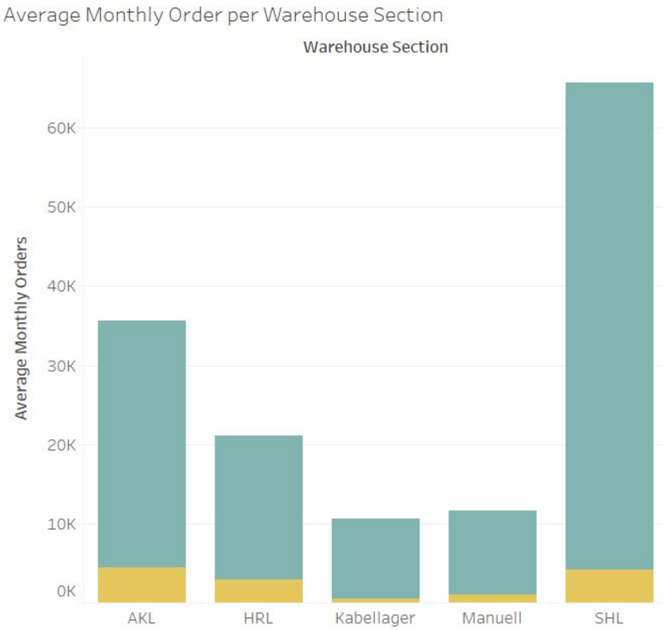
245.410
Avg. Orders per
Quarter

1.023.468
Avg. Orders per
Year
excluding 2011 & 2020 due to incomplete data

2.2.2 KPI Number of Orders per Origin





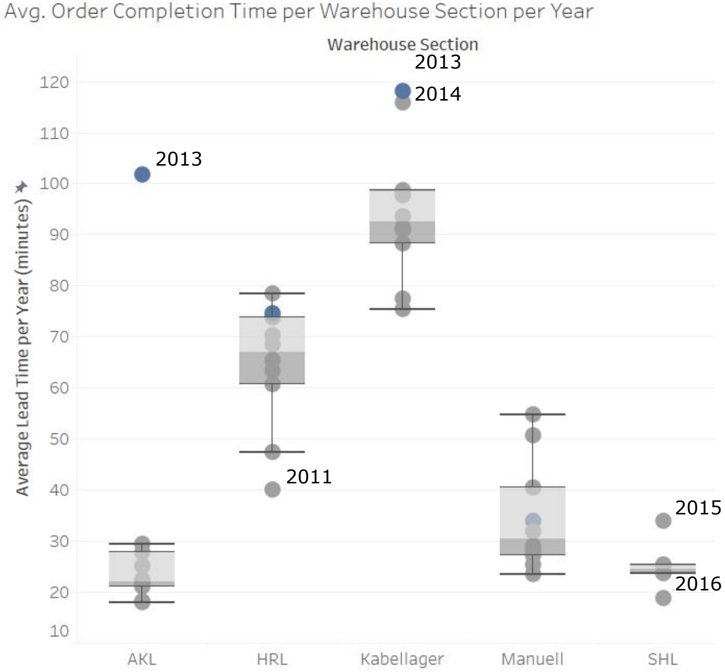


Insights

- Majority of orders were made by customers, making up **89,2%** while store orders constituted **10,8%**. While recording a higher share of pick volume, customer orders retained majority.
- **SHL** fulfilled the highest customer orders on average, followed by **AKL**.
- On average. **AKL** fulfilled the highest number of store orders, closely followed by **SHL**.

2.2.3 Order Completion Time

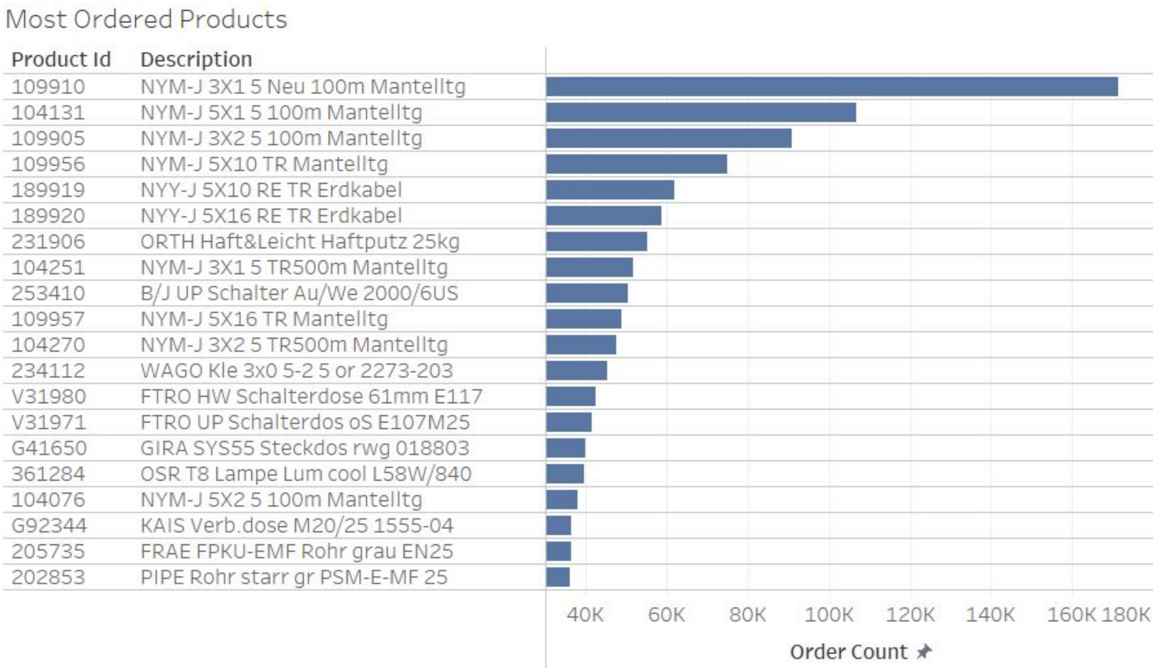
Warehouse Section	Avg. Lead Time (mins)
SHL	23,6
Manuell	34,8
AKL	37,7
HRL	64,7
Kabellager	93,2



Insights

- **SHL** achieved the fastest lead time on average; it took **23,6 minutes** to fulfill order in that section on average.
- **Kabellager** was the slowest in pick time for orders, taking an average of **93,2 minutes**.

2.2.4 Most ordered products overall



Insights

- The top ordered product overshadowed the rest at **171,4k** ordered throughout.
- This was **192%** higher than the average of the top 20 ordered products.
- Some product IDs refer to different versions of the same product. However, we are handling them as unique products.

3. Summary and Next Steps

3.1 Summary of Findings

Key Findings	
1. Scale & Types	Extensive dataset with 33.888.990 rows with mix of data types, non-numeric & repeating order numbers, missing-, zero- & negative values, duplicates, outliers and special characters
2. Cleaning Impact	After quality checks the dataset was reduced to 33.277.547 rows
3. KPI Selection & Statistical Measures	1. SHL's Significant Impact on Pick Volume 2. Seasonal Fluctuations in Pick Volume 3. Dominance of Specific Product Groups and Sections 4. Warehouse Performance Variations
4. Missing Data Description	1. No differentiation of quantity units. Leads to inaccurate evaluation of the pick numbers as 1 “St” is not equal to 1 “Pa” 2. Updated products have the same name as older versions

3.2 Next Steps

Next Steps
1. Evaluating Sections of Efficiency
2. Focus on Underperforming Product Groups
3. Evaluating Seasonal Demand (Fluctuation in pick volumes)
4. Application to the 3-week absence of the manager
5. Customize the dashboard for an immediate understanding of the status
6. Definition of thresholds (Data Quality, Performance, Errors, Anomaly Detection)
7. Creating dashboards for forecasting business & better decision making

Obeta Project

3. Dashboard Presentation and Business Analysis

Ankit Gupta (**590516**) *Ankit.Gupta@Student.HTW-Berlin.de*
Bhoomika Jagadeesha (**590573**) *Bhoomika.Jagadeesha@student.htw-berlin.de*
Majeed Abdul-Razak (**590507**) *Majeed.Abdul-Razak@Student.HTW-Berlin.de*
Rushikesh Barge (**590495**) *Rushikesh.Barge@Student.HTW-Berlin.de*
Pei Hua Yung (**590659**) *Pei.Yung@Student.HTW-Berlin.de*

19 Jan 2024

Hochschule für Technik
und Wirtschaft Berlin
University of Applied Sciences

1.PICK — 2.ORDER — 3.EFFICIENCY — 4.YEAR 2020

1.PICK

- Picks per Warehouse
- Picks per Year
- Picks per Product Group
- Picks per Origin

2.ORDER

- Warehouse Orders per Year
- Orders per Warehouse Section
- Orders per Product Group
- Orders per Origin

3.EFFICIENCY

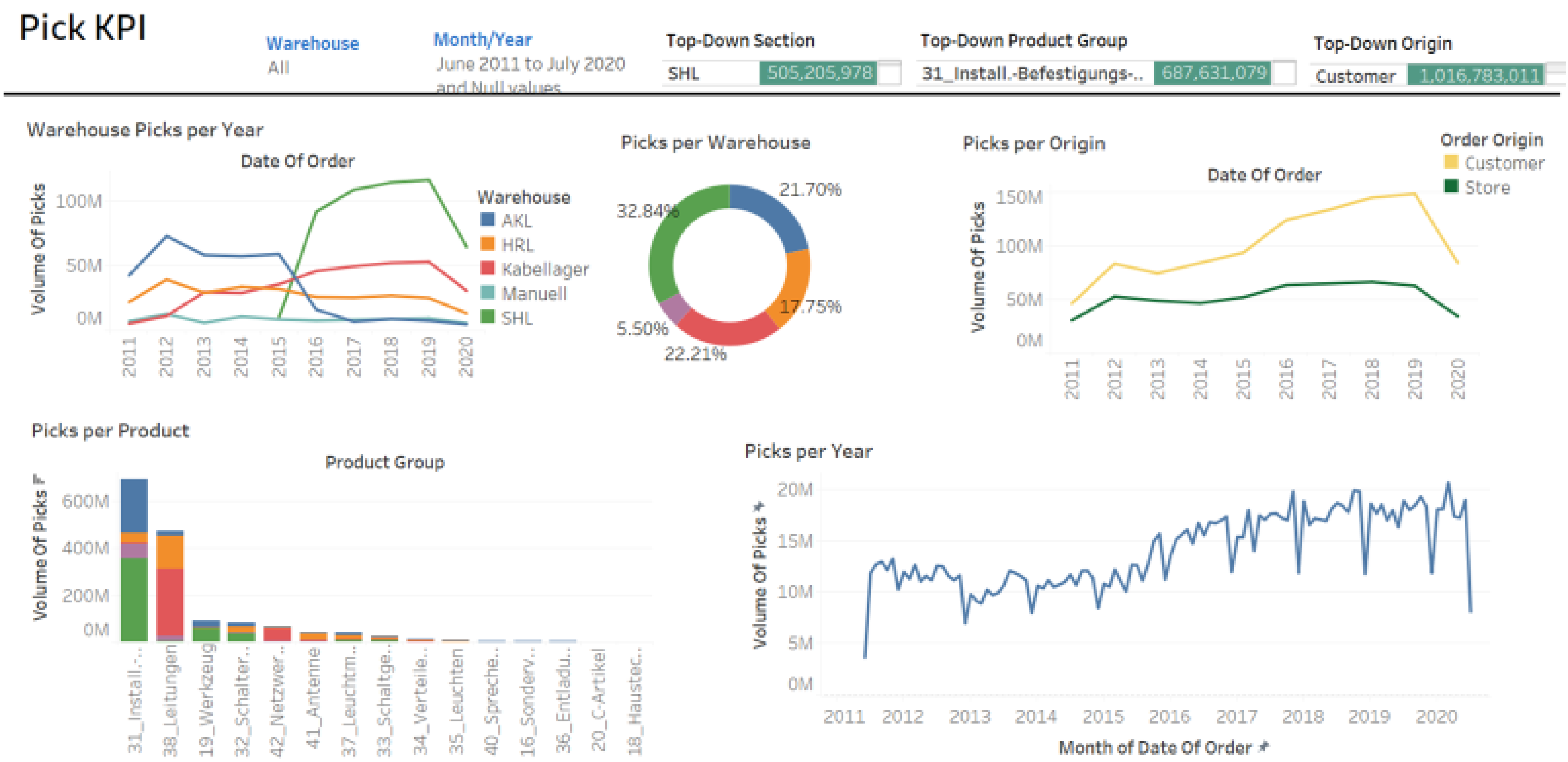
- Yearly Average Pick Efficiency per Order by Warehouse
- Yearly Average Pick Efficiency per Product by Warehouse

4.YEAR 2020

- Total of Picks per Warehouse, year of 2020
- Percentage Change per week
- Percentage Change per month

1.Picks

- Picks per Warehouse
- Picks per Origin
- Picks per Product Group
- Picks per Year

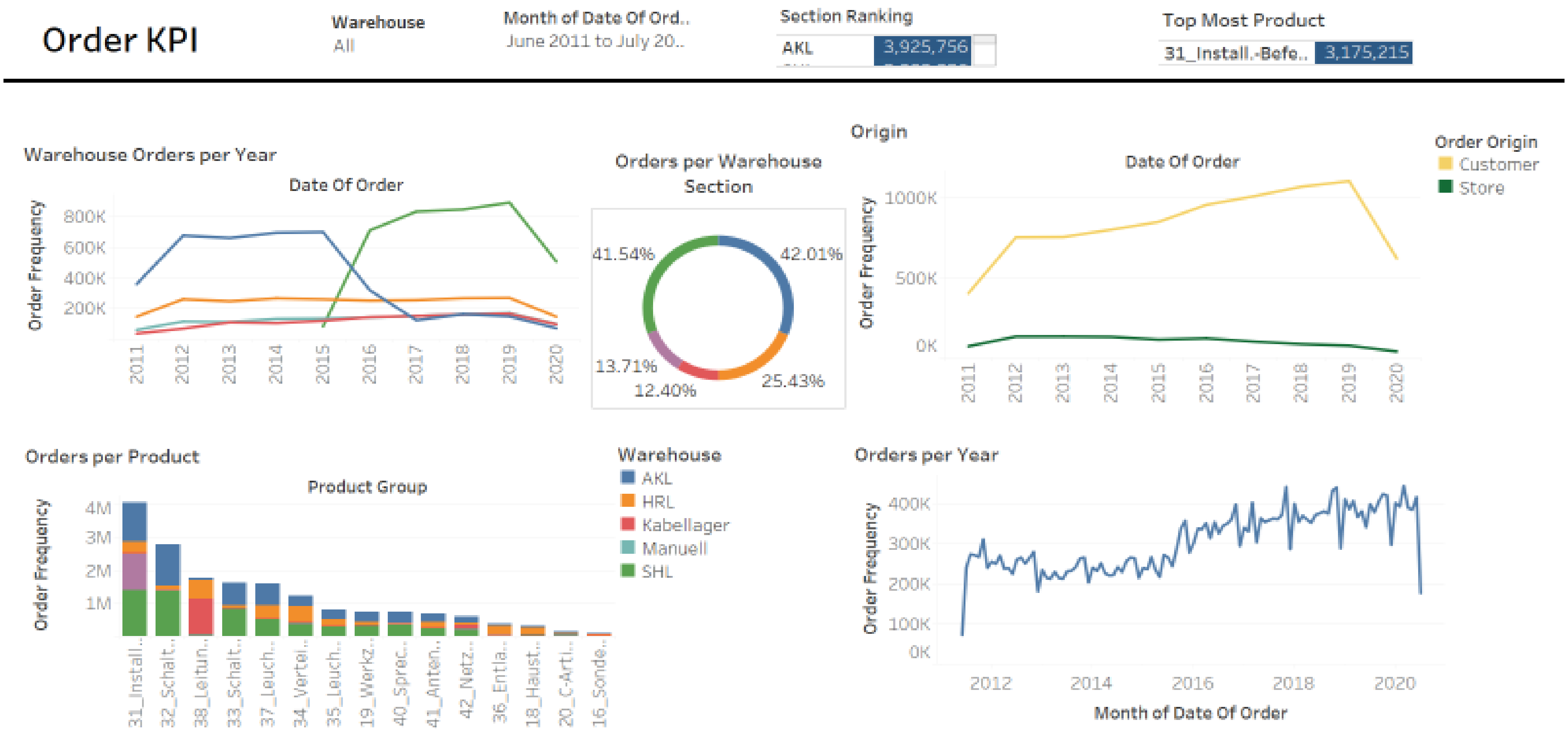

Pick KPI
Warehouse
All
Month/Year
June 2011 to July 2020 and Null values
Top-Down Section
SHL 505,205,978
Top-Down Product Group
31_Install.-Befestigungs-.. 687,631,079
Top-Down Origin
Customer 1,016,783,011
Warehouse Picks per Year
Date Of Order
Volume Of Picks
100M
50M
0M
2011 2012 2013 2014 2015 2016 2017 2018 2019 2020
Warehouse
AKL
HRL
Kabellager
Manuell
SHL
Picks per Warehouse
21.70%
32.84%
17.75%
5.50%
22.21%
Picks per Origin
Date Of Order
Order Origin
Customer
Store
Volume Of Picks
150M
100M
50M
0M
2011 2012 2013 2014 2015 2016 2017 2018 2019 2020
Picks per Product
Product Group
Volume Of Picks
600M
400M
200M
0M
31_Install.-..
38_Leitungen
19_Werkzeug
32_Schalter..
42_Netzwer..
41_Antenne
37_Leuchtm..
33_Schaltge..
34_Verteile..
35_Leuchten
40_Spreche..
16_Sonderv..
36_Entladu..
20_C-Artikel
18_Haustec..
Picks per Year
Volume Of Picks
20M
15M
10M
5M
0M
2011 2012 2013 2014 2015 2016 2017 2018 2019 2020
Month of Date Of Order

Insights

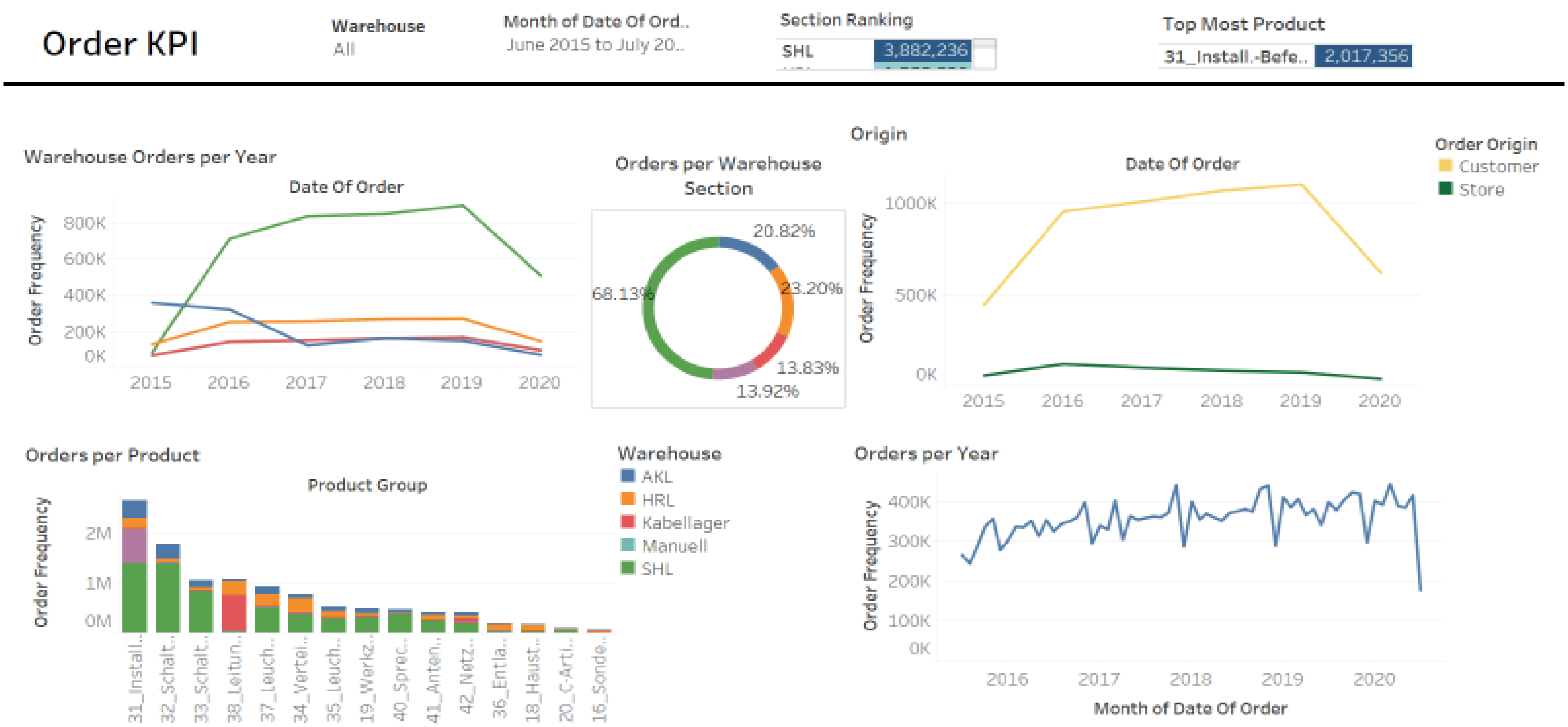
- **SHL** has maximum Volumes of Pick from 2015 onwards
- **Customer channel** has the most number of picks
- The pick volumes show **upward trend in March** and **downward trend in December**
- Product Group **31_Install_Bestifigung** has the highest number of picks

2. Orders

- Warehouse Orders per Year
- Orders per Warehouse Section
- Orders per Product
- Orders per Origin



Order KPI
Warehouse
All
Month of Date Of Ord..
June 2011 to July 20..
Section Ranking
AKL
3,925,756
Top Most Product
31_Install.-Befe..
3,175,215
Warehouse Orders per Year
Date Of Order
Order Frequency
800K
600K
400K
200K
2011
2012
2013
2014
2015
2016
2017
2018
2019
2020
Orders per Warehouse Section
41.54%
42.01%
13.71%
12.40%
25.43%
Origin
Date Of Order
Order Frequency
1000K
500K
0K
Order Origin
Customer
Store
Orders per Product
Product Group
Order Frequency
4M
3M
2M
1M
31_Install..
32_Schalt..
38_Leitun..
33_Schalt..
37_Leuch..
34_Verteil..
35_Leuch..
19_Werkz..
40_Sprec..
41_Anten..
42_Netz..
36_Entla..
18_Haust..
20_C-Arti..
16_Sonde..
Warehouse
AKL
HRL
Kabellager
Manuell
SHL
Orders per Year
Order Frequency
400K
300K
200K
100K
0K
2012
2014
2016
2018
2020
Month of Date Of Order


Order KPI
Warehouse
All
Month of Date Of Ord..
June 2015 to July 20..
Section Ranking
SHL 3,882,236
Top Most Product
31_Install.-Befe.. 2,017,356
Warehouse Orders per Year
Date Of Order
Order Frequency
800K
600K
400K
200K
0K
2015
2016
2017
2018
2019
2020
Orders per Warehouse Section
20.82%
23.20%
68.13%
13.83%
13.92%
Origin
Date Of Order
Order Frequency
1000K
500K
0K
Order Origin
Customer
Store
Orders per Product
Product Group
Order Frequency
2M
1M
0M
31_Install..
32_Schalt..
33_Schalt..
38_Leitun..
37_Leuch..
34_Vertei..
35_Leuch..
19_Werkz..
40_Sprec..
41_Anten..
42_Netz..
36_Entla..
18_Haust..
20_C-Arti..
16_Sonde..
Warehouse
AKL
HRL
Kabellager
Manuell
SHL
Orders per Year
Order Frequency
400K
300K
200K
100K
0K
Month of Date Of Order

Insights

- **SHL** process about 70% of the total orders for the period 2015 onwards
- **Customer channel** has the most number of orders
- The number of orders show **downward trend in December every year**
- Product Group **31_Install_Bestifigung** has the highest number of orders

3. Pick Efficiency

- Yearly average pick efficiency per order by Warehouse section
- Yearly average pick efficiency per product by Warehouse section

Pick Efficiency

Insights: Before 2015 Manuell and AKL were most efficient, though AKL had a sharp decrease and increase in 2013

Pick Efficiency

Insights: From 2015 SHL & AKL are the most efficient, Manuell had a sharp decline after 2017

Pick Efficiency

SELECT Warehouse
Multiple values

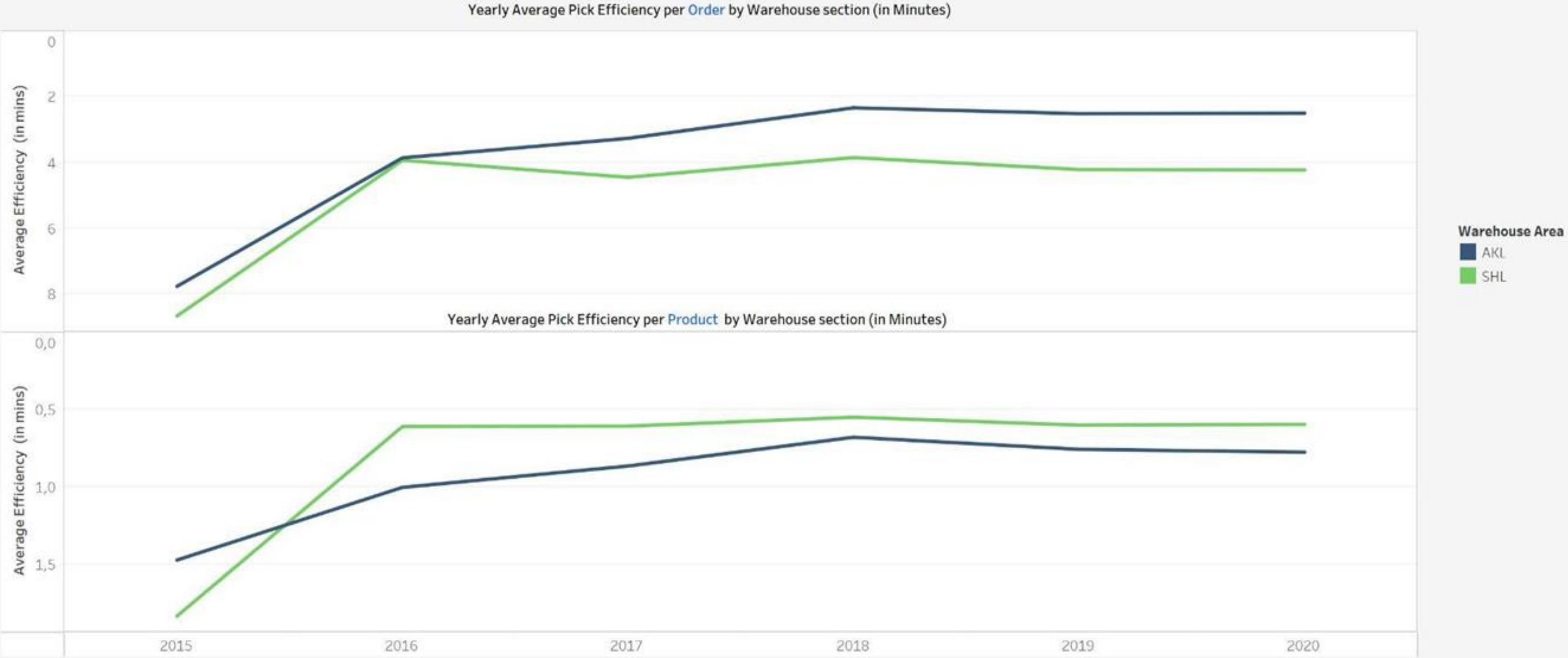
SELECT Range
2014 to 2020
and Null values

Most efficient Section per Order

SHL	4,23 Mins/Order

Most efficient Section per Product

SHL	0,62 Mins/product



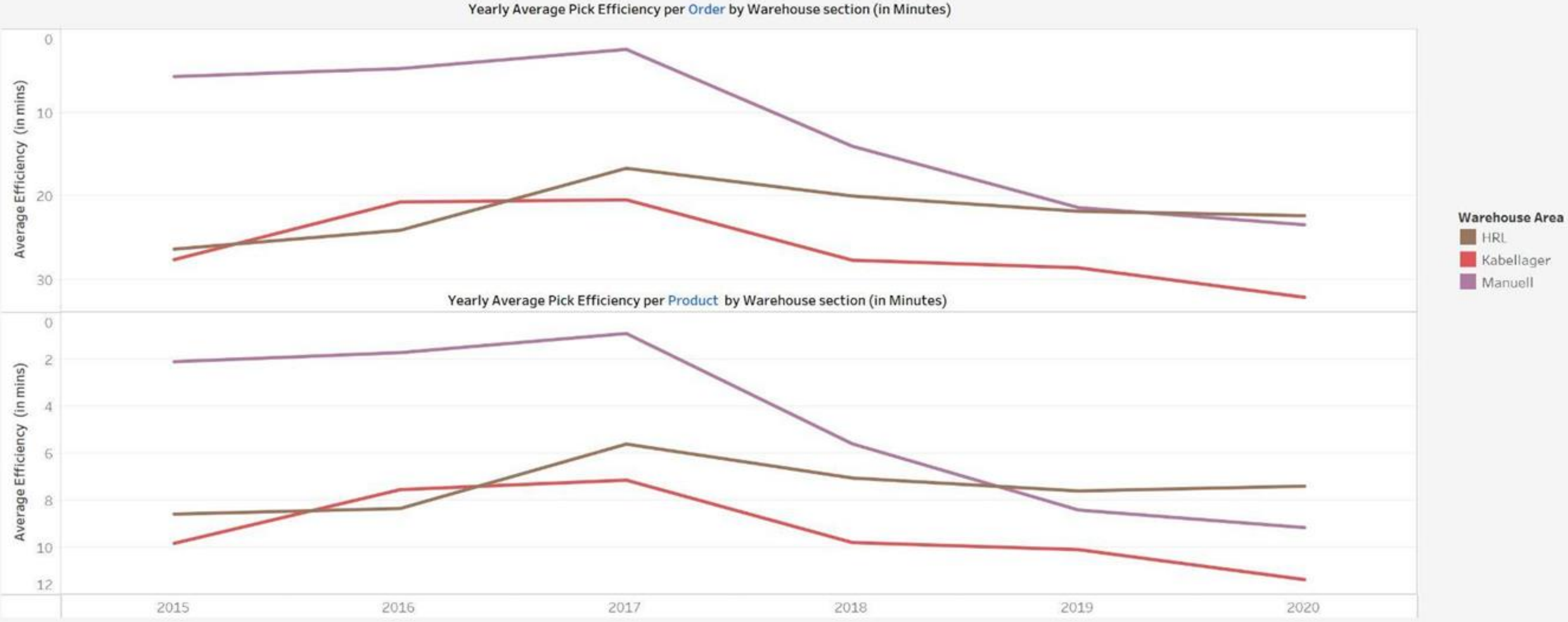
Pick Efficiency

SELECT Warehouse
Multiple values

SELECT Range
2014 to 2020
and Null values

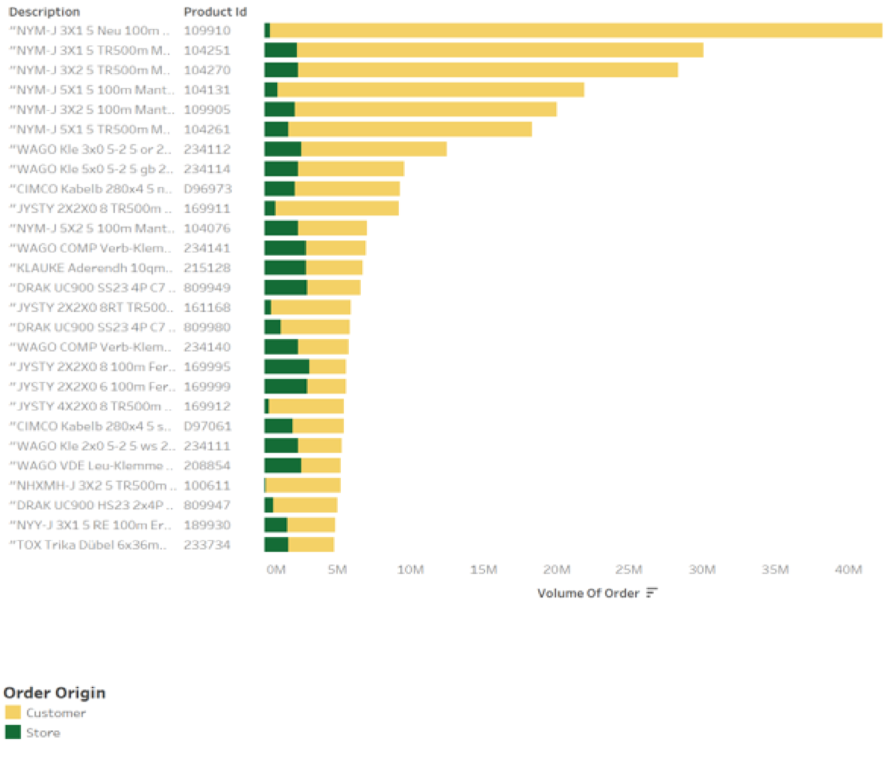
Most efficient Section per Order	
SHL	4,23 Mins/Order

Most efficient Section per Product	
SHL	0,62 Mins/product

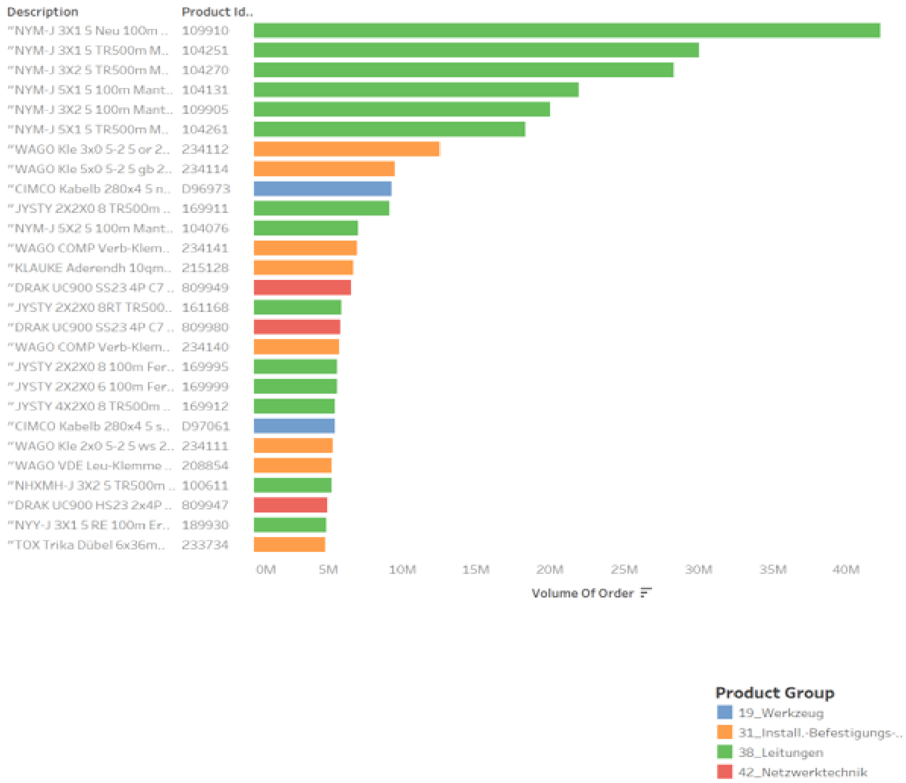


Insights: HRL, Kabellager and Mauell are at optimum efficiency in 2017 and all continue to declined afterwards

Majority of top 30 products' orders come directly from customer



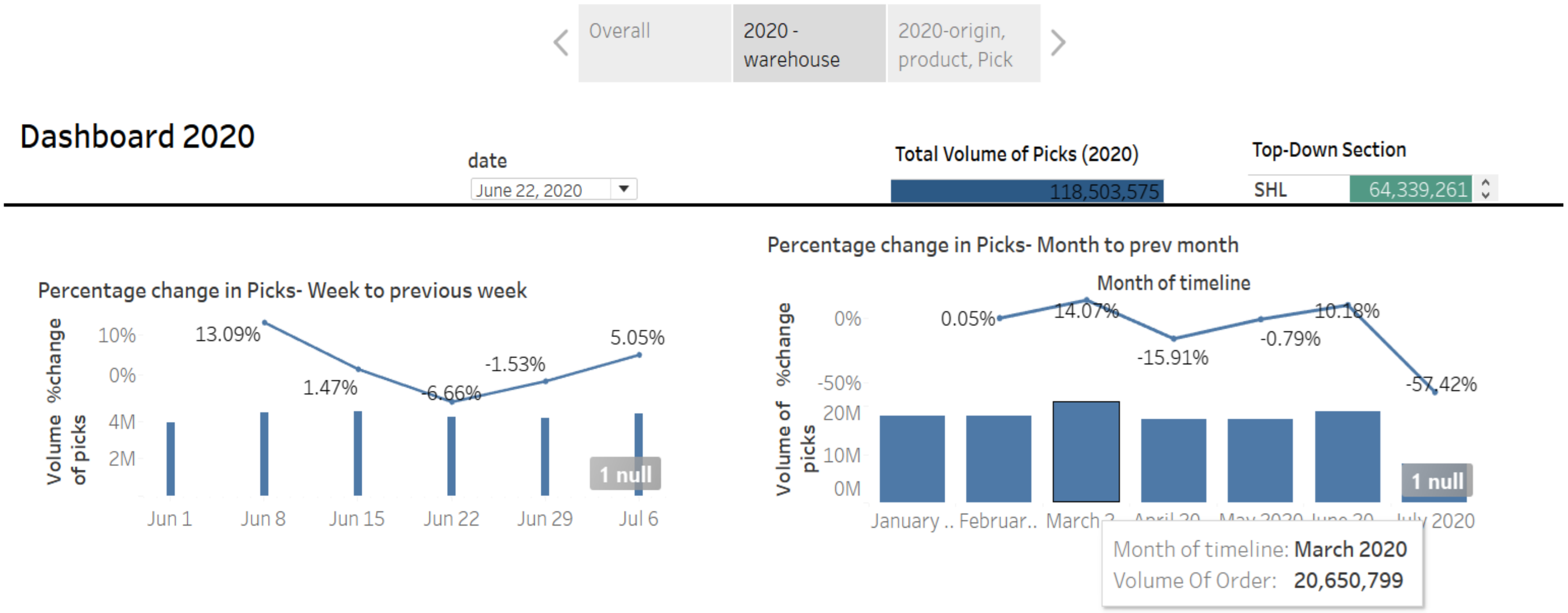
Top 30 products belong to 4 product groups, they are mostly handled by AKL and SHL



4. Year 2020

- Total of Picks per Warehouse, year of 2020
- Percentage Change per week
- Percentage Change per month

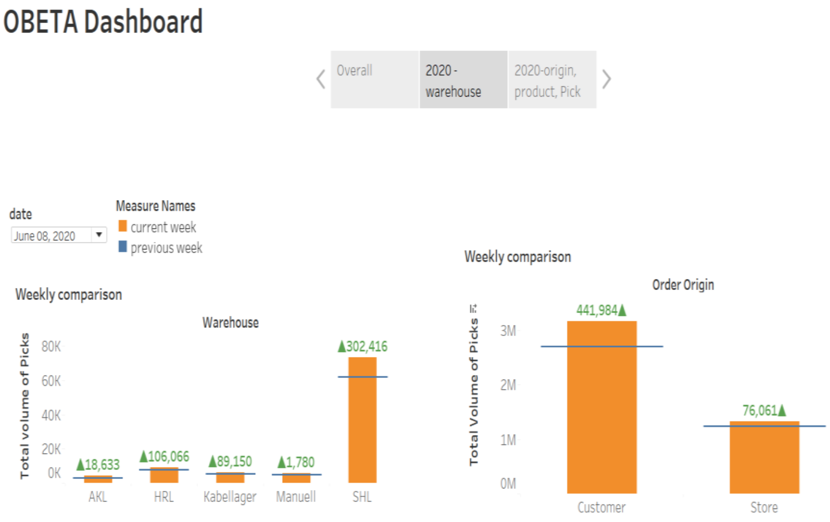
OBETA Dashboard



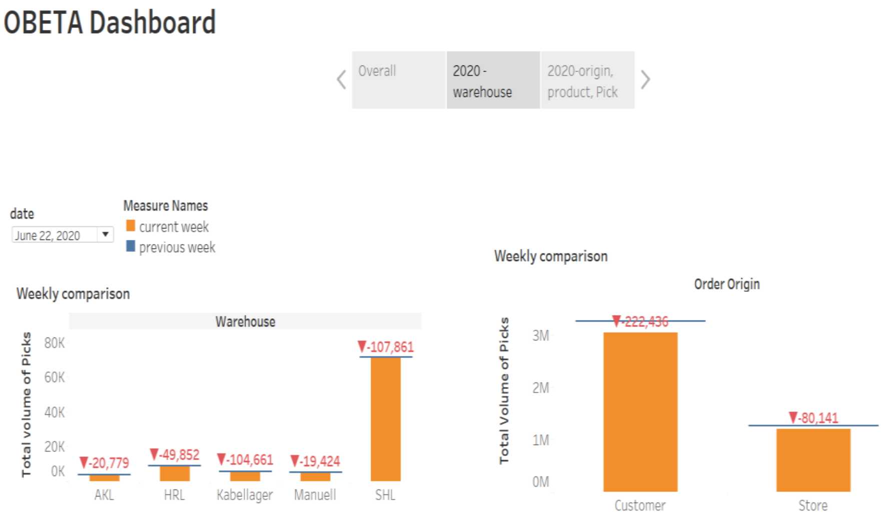
Insights: 1. **Peak week:** June 2nd week - 4,540,308 units; **Peak month**: March 2020 - 20,650,799 units.

2. The percentage of growth is on a declining trend over time.

Weekly Analysis



Peak Week (June 2nd week)



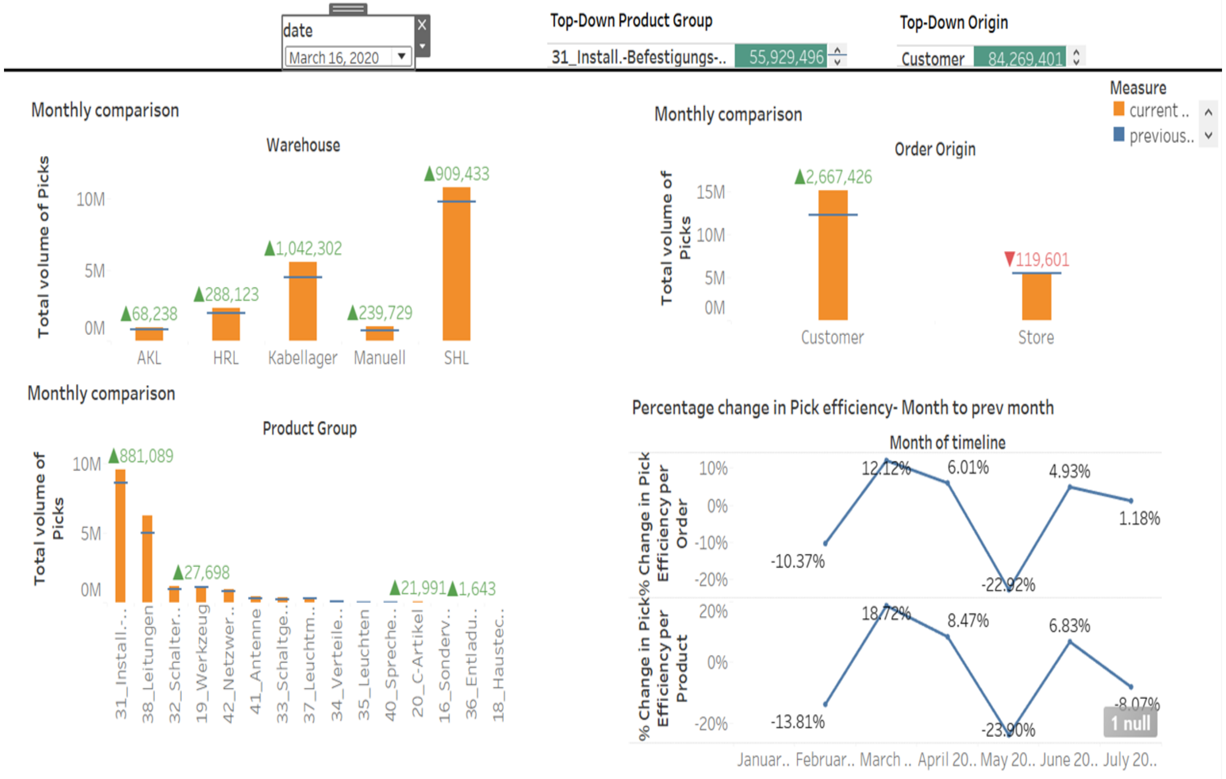
Trough Week (June 3rd Week)

Insights:

1. 13.73% increase in **SHL** pick volumes.
2. Notable surge - 16.3% in **customer** orders.

1. Slight downturn of 1.26% in the **SHL** warehouse despite being the busiest.
2. Notable reduction of 6.81% in **customer/online orders**.

Monthly Analysis



Peak Month (March)

Insights:

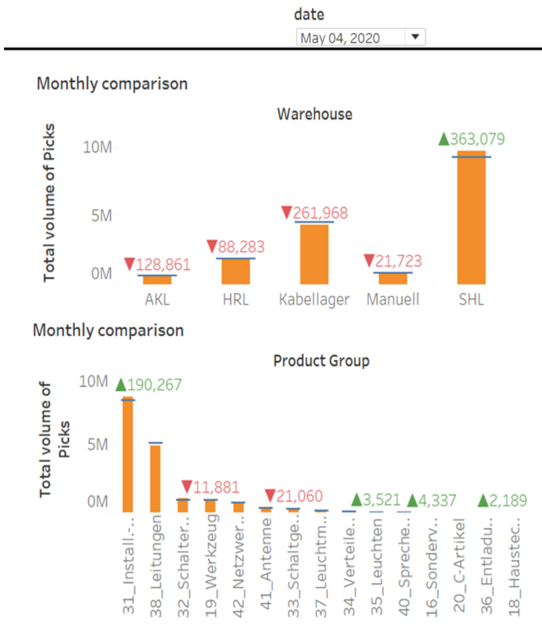
1. Robust Growth- 22.98% increase in pick volumes for **Kabellager**.

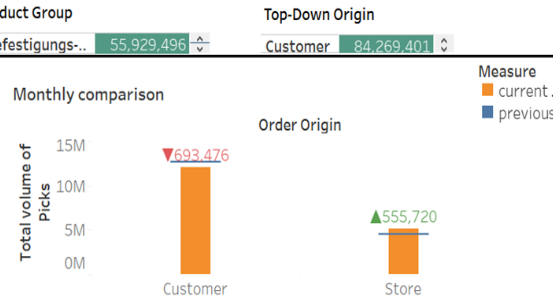
2. Followed by **SHL** by 9.2%.

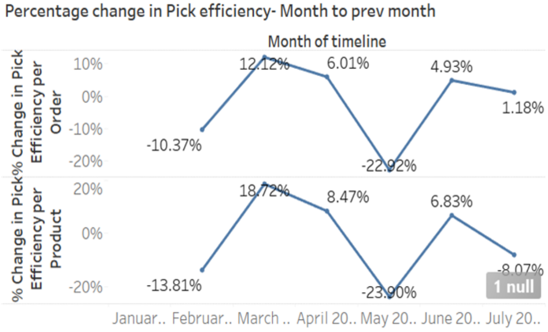
3. While SHL has reached its optimum picks with no significant recent growth, Kabellager has experienced considerable expansion on the other hand.

4. 21.42% increase in **Customer** channel.

5. Clearly, the month at its peak aligns with the substantial percentage increase in pick efficiency.







Trough Month (May)

Insights:

1. 5.7% decline in pick volumes was noted in **Kabellager.**

2. In contrast, **SHL** section experienced a positive trend with increase in pick volumes by 3.92%.

3. Customer/Online orders: decline of 5.44%;
   Store orders: rise of 12.04%.

4. Trough month coincides with lowest percentage increase in pick volumes.

5.0 Summary

1. **SHL** and **AKL** are the most efficient per product and order;

   HRL, **Kabellager** and **Manuell** are in decline of efficiency since 2017.
2. The growth curve is flattening out.
3. SHL being the busiest warehouse has reached its optimum. It reflects a less substantial increase in pick volumes.

   Slight growth could be observed in Kabellager.
4. The highest activity in 2020 occurred during the 2nd week of June and peak month (both in volume and efficiency) was noted in March.
5. There is a seasonal pattern in the growth. The peak occurs in October/November and the low point occurs in December every year.

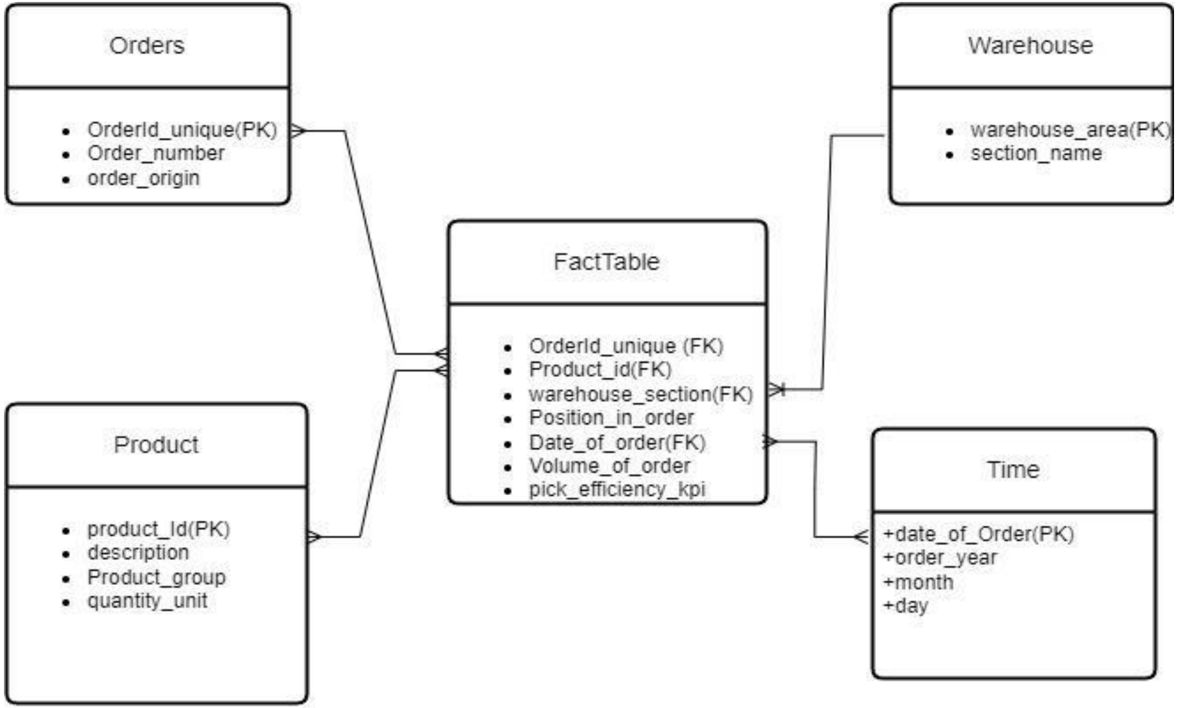
Recommendation & Best Practice

- Upgrade Technology:
    - Investment in automated processes (SHL, AKL)
    - Investment in management softwares
    - Investment in real-time inventory tracking and monitoring system
    - Investment in website and app refinement

- Forecast Seasonal Demand:
    - Staffing adjustment for higher demand quarters
    - Future inventory predictions based on historical data

- Storage Space Optimization:
    - More space for products with high demand
    - Less space for products with low demand

Appendix: ER Model



Hochschule für Technik
und Wirtschaft Berlin

University of Applied Sciences

www.htw-berlin.de