Funnel Analysis

Presented by: Eunike Pardede
Last Updated: June 27, 2023

Introduction

XYZ logistic company is responsible for developing logistics facilities tailored to the unique requirements of their diverse clients.

As a branch company of XYZ, ABC must deliver space for a logistics solution. In this case, ABC reappropriated existing spaces in a building to accommodate the client.

Objective

As a data analyst, we have to support ABC colleagues in identifying opportunities for XYZ Supply Chain.

The goal is to **locate new clients** in existing buildings on an ongoing basis by giving recommendations of suitable spaces for clients **based on their preferences and requirements.**

User Inputs

1. Existing building information
2. Client preferences

Output

A list of recommended spaces with detailed information, including its location, size, cost, etc.

Filters

1. Location
2. Industry
3. Space Availability
4. Annual Lease Rate/sqm
5. Starting Date

Data sets

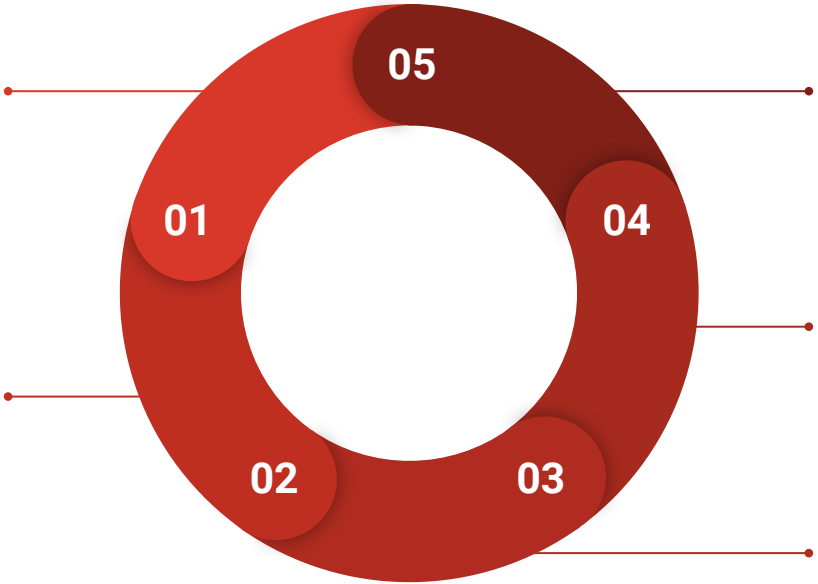
Gather data on existing buildings, including their characteristics and location, lease, and existing contracts.

facilities	leases	contracts
- Facility ID - Coordinates - City - SQM - Employees - BREEAM Certification	- Facility ID - Lease ID - Lease End Date - SQM - Lease Rate (EURk)	- Customer - Industry - Contract End Date - Revenue (EURk) - Gross Profit (EURk)

STEPS

Data Sourcing & Inspection
- Gather data on existing buildings, lease, & existing clients.
- Identify common columns

Data Cleaning
- Rename columns name
- Create new columns
- Fix the datatype
- Check null items



Visualization
Develop interactive dashboard

Analysis
- Perform occupancy analysis
- Analyze the availability of suitable spaces
- Filter out any spaces that do not meet the client's requirements.

Data Integration & Preparation
- Merge 3 data sets into 1 dataframe
- Create calculated fields

Tools:

libraries: pandas
Jupyter Notebook

Data Sourcing & Inspection

Import necessary libraries & load the dataset ¶

```
#import library
import pandas as pd
from pathlib import Path
```

```
#load the dataset
file = Path('/[redacted].xlsx')

facilities = pd.read_excel(file,sheet_name='Facilities')
leases= pd.read_excel(file,sheet_name='Leases')
contracts = pd.read_excel(file,sheet_name='Contracts')
```

Data Sourcing & Inspection

1. facilities dataset

```
#inspect
facilities.head()
```

	Facility ID	Coordinates	City	SQM	Employees	BREEAM Certification
0	FCLTY001	60.1733, 24.9414	Helsinki	13830.427441	50	Yes
1	FCLTY002	55.6759, 12.5655	Copenhagen	10286.810846	40	No
2	FCLTY003	59.9142, 10.7522	Oslo	14606.531554	120	Yes
3	FCLTY004	52.3731, 4.8922	Amsterdam	26400.231922	75	Yes
4	FCLTY005	51.2277, 6.7735	Düsseldorf	10797.883796	50	No

Coordinates column needs to be split into Latitude & Longitude

```
len(facilities.City.unique())
```

```
25
```

```
len(facilities['Facility ID'].unique())
```

```
43
```

There are **43 facilities** that are located in **25 cities**

Data Sourcing & Inspection

2. leases dataset

```
leases.head()
```

	Facility ID	Lease ID	Lease End Date	SQM	Lease Rate (EURk)
0	FCLTY001	SEFMO2	2022-12-29	3130	38
1	FCLTY002	1FR2AY	2026-10-11	2108	82
2	FCLTY002	9MSL7A	2024-03-12	2880	113
3	FCLTY002	40UZJ9	2030-11-21	873	22
4	FCLTY003	EQGOXS	2025-10-22	10361	114

```
len(leases['Lease ID'].unique())
```

```
92
```

33 of 92 leases have expired.

There are multiple leases within 1 facility.
For example, FCLTY002 has 3 leases.

Data Sourcing & Inspection

3. contracts dataset

```
contracts.head()
```

	Customer	Industry	Contract End Date	Revenue (EURk)	Gross Profit (EURk)
0		Technology	2022-12-29	8610.34	3720.40
1		Telecommunications	2026-10-11	9742.55	4058.36
2		Telecommunications	2024-03-12	9473.69	4153.19
3		Telecommunications	2030-11-21	4481.33	1202.17
4		Banking	2025-10-22	3390.76	771.45

```
len(contracts['Customer'].unique())
```

```
89
```

89 contracts were from different clients, of which 31 contracts have ended.

Data Cleaning

Facilities Dataset
- Add new columns: Latitude & Longitude
- Delete Coordinates column.
- Rename SQM to FacilitySQM.
- Fix the data type of FacilitySQM from float to integer.

Leases & Contracts Dataset
- Rename Contract End Date column in '*leases*' &
- Rename Lease End Date column in '*contracts*' to End Date

```
#Split latitude and longitude from Coordinates columns
facilities[['Latitude','Longitude']] = facilities['Coordinates'].str.split(",",expand=True)
del facilities['Coordinates']
```

```
#rename 'SQM' columns in 'facilities' dataset to be 'FacilitySQM'
facilities.rename(columns={'SQM':'FacilitySQM'},inplace=True)
```

```
#change datatype from float to int
facilities['FacilitySQM'] = facilities['FacilitySQM'].astype(int)
```

```
#rename 'Contract End Date' columns in 'leases'& 'Lease End Date' columns in 'contracts' to be 'End Date'
leases.rename(columns={'Lease End Date': 'End Date','SQM':'LeaseSQM'}, inplace=True)
contracts.rename(columns={'Contract End Date': 'End Date'}, inplace=True)
```

Data Cleaning

Check null items

```
all_data.isnull().sum()
```

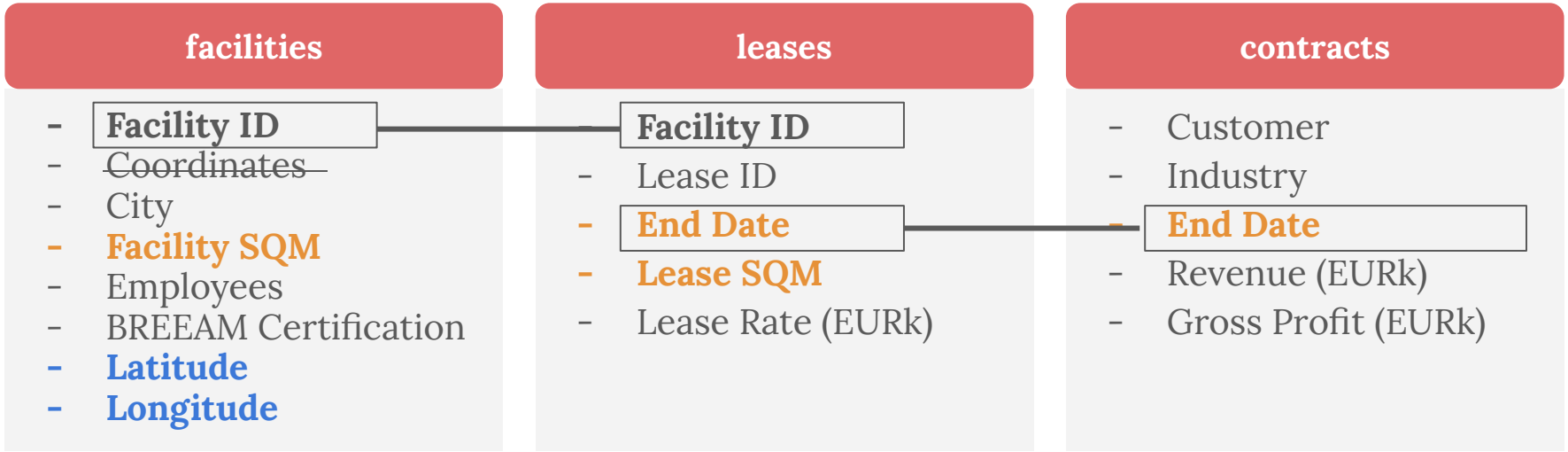
```
Facility ID             0
City                    0
FacilitySQM             0
Employees               0
BREEAM Certification    0
Latitude                0
Longitude               0
Lease ID                0
End Date                0
LeaseSQM                0
Lease Rate (EURk)       0
Customer                3
Industry                3
Revenue (EURk)          3
Gross Profit (EURk)     3
Availability            0
Occupancy SQM           0
Lease_rate_per_sqm      0
dtype: int64
```

There are 3 leases whose previous tenants are not identified.

```
all_data.loc[all_data.Customer.isnull()]
```

	Facility ID	City	FacilitySQM	Employees	BREEAM Certification	Latitude	Longitude	Lease ID	End Date	LeaseSQM	Lease Rate (EURk)	Customer	Industry	Revenue (EURk)	Gross Profit (EURk)
36	FCLTY017	Łódź	16278	100	Yes	51.9194	19.1451	TM6WCS	2025-12-15	9618	92	NaN	NaN	NaN	NaN
58	FCLTY026	Antwerp	11291	80	Yes	51.2217	4.3997	I9ZLYW	2022-01-13	9357	35	NaN	NaN	NaN	NaN
84	FCLTY040	Helsinki	18698	75	No	60.1649	24.9486	P5M7J0	2021-06-22	10249	72	NaN	NaN	NaN	NaN

Data Integration



Data Integration

Merge 3 data sets into a data frame & save the data frame into Excel file.

```
#Combine visits and cart using a left merge
building_info = pd.merge(facilities,leases,how ='left')
building_info.head()
```

```
all_data=building_info.merge(contracts,how='left')
all_data.head()
```

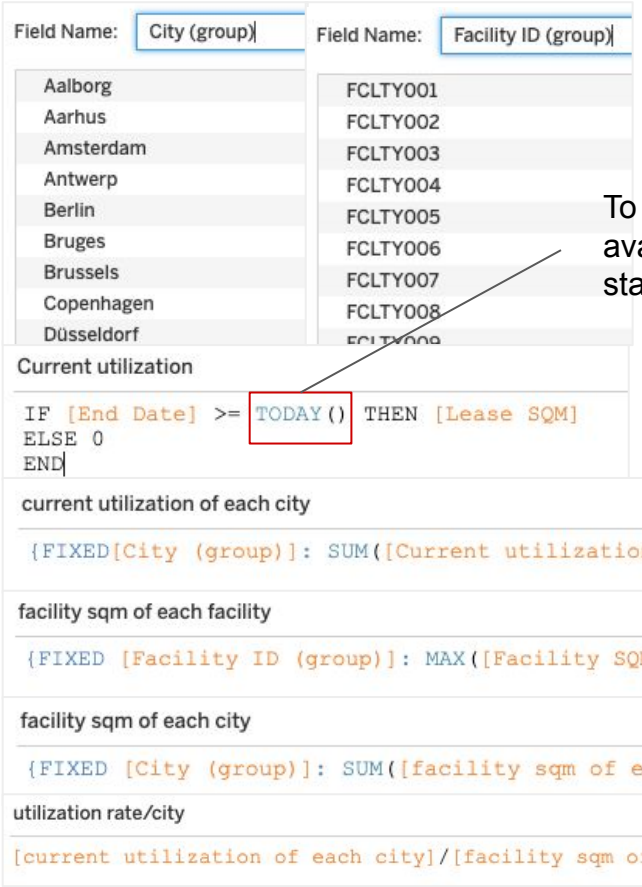
	Facility ID	City	FacilitySQM	Employees	BREEAM Certification	Latitude	Longitude	Lease ID	End Date	LeaseSQM	Lease Rate (EURk)	Customer	Industry	Reve (EU
0	FCLTY001	Helsinki	13830	50	Yes	60.1733	24.9414	SEFMO2	2022-12-29	3130	38		Technology	861(
1	FCLTY002	Copenhagen	10286	40	No	55.6759	12.5655	1FR2AY	2026-10-11	2108	82		Telecommunications	974
2	FCLTY002	Copenhagen	10286	40	No	55.6759	12.5655	9MSL7A	2024-03-12	2880	113		Telecommunications	947
3	FCLTY002	Copenhagen	10286	40	No	55.6759	12.5655	40UZJ9	2030-11-21	873	22		Telecommunications	448
4	FCLTY003	Oslo	14606	120	Yes	59.9142	10.7522	EQGOXS	2025-10-22	10361	114		Banking	339(

```
all_data.to_excel("all_data.xlsx")
```

Data Preparation

Create some calculated fields:

- Utilization rate for each city based on today's date.
- Available spaces per facility
- Average lease rate/sqm (EUR)
- Filter



Data Preparation

Create some calculated fields:

- Utilization rate for each city based on today's date.
- Available spaces per facility
- Average lease rate/sqm (EUR)
- Filter

A parameter to identify the date the new prospective client will lease the space

Utilization based on start date

```
IF [End Date] >= [Start Date] THEN [Lease SQM]
ELSE 0
END
```

available space per facility

```
Max([Facility SQM])
-Sum([Utilization based on start date])
```

Data Preparation

Create some calculated fields:

- Utilization rate for each city based on today's date.
- Available spaces per facility
- Average lease rate/sqm (EUR)
- Filter

lease rate/sqm(EUR)

```
[Lease Rate (EURk)]/[Lease SQM]
```

Avg. lease rate/sqm (EUR)

```
{FIXED [Facility ID (group)]:
AVG([lease rate/sqm(EUR)])*1000}
```

Data Preparation

Create some calculated fields:

Utilization rate for each city based on today's date.

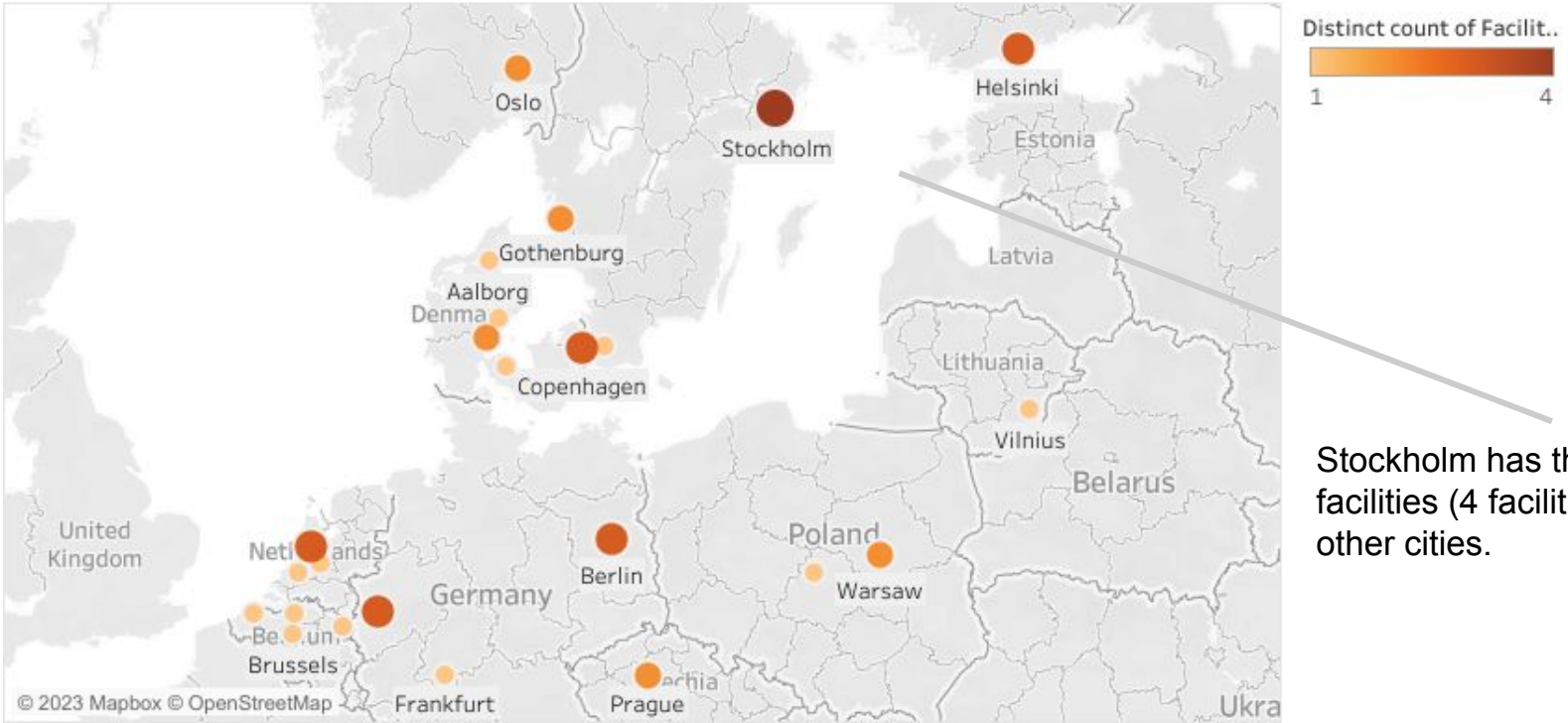
Available spaces per facility

Average lease rate/sqm (EUR)

Filter

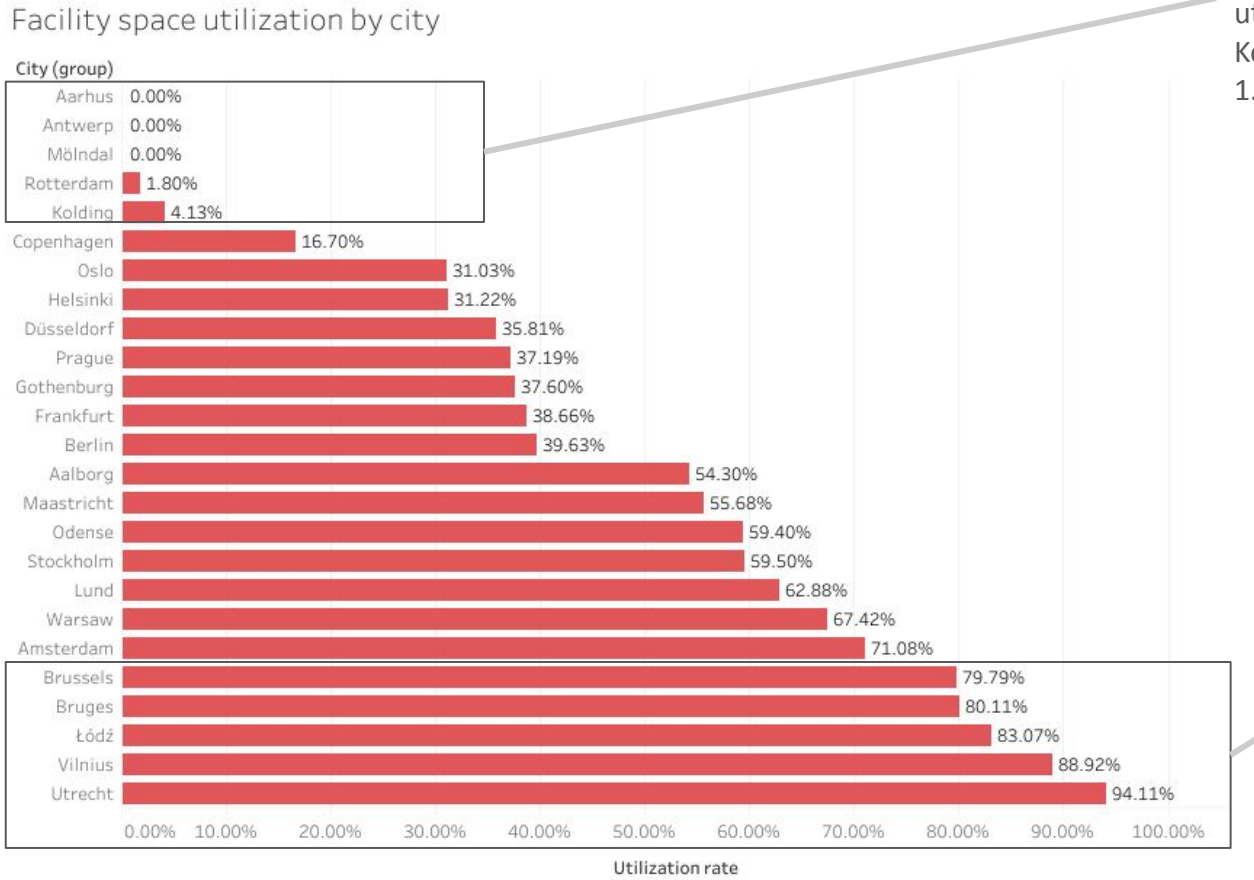
Data Analysis

Location of Facilities



Stockholm has the most facilities (4 facilities) among other cities.

Data Analysis

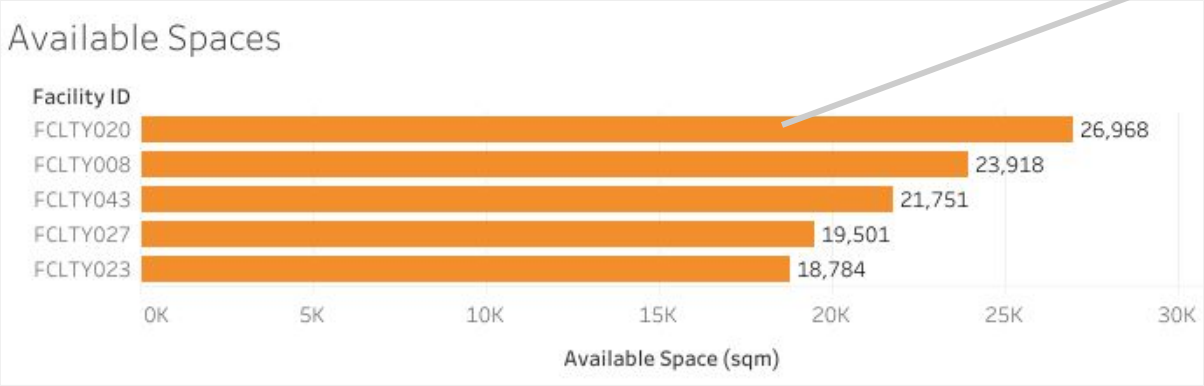


Aarhus, Antwerp, Mölndal have zero utilization rates. Rotterdam and Kolding have lower utilization rates, by 1.8% and 4.13%.

Top 5 cities that have higher occupancy rates. Facilities in Utrecht have been utilized by 94.11%

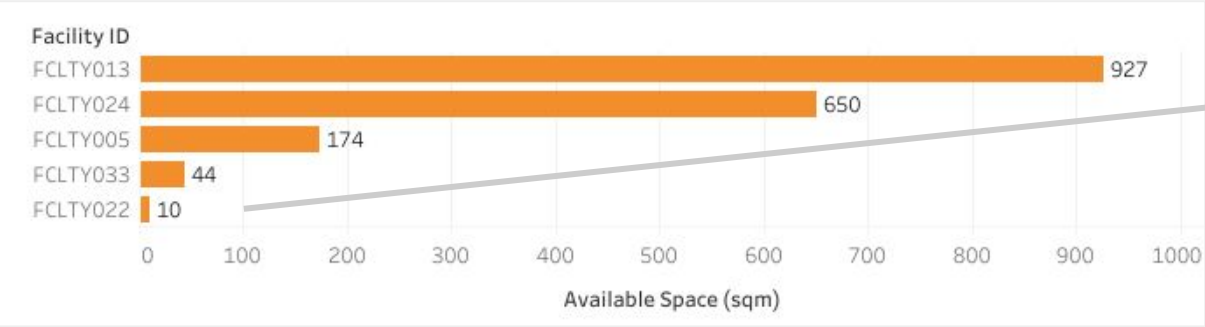
Data Analysis

TOP 5



Currently, FCLTY020 (in Mölndal) is not occupied by any clients.

BOTTOM 5



Meanwhile, FCLTY022 (in Stockholm) only has an area of 10 square meters left.

Data Analysis

Summary of facilities

- City
- Industry
- Average Lease rate/sqm
- Availability of BREEAM certificate
- Number of employees
- Utilization status

Details

City	Industry	Facility ID	Avg. lease r..	BREEAM	Employees.	
Aalborg	Construction	FCLTY010	€10.35	No	40	6,173
Aarhus	Technology	FCLTY015	€10.28	Yes	60	0
Amsterdam	Construction	FCLTY004	€8.93	Yes	75	12,037
		FCLTY036	€4.71	Yes	75	15,847
	Healthcare	FCLTY024	€204.98	Yes	150	25,872
Antwerp	Null	FCLTY026	€3.74	Yes	80	0
Berlin	Energy	FCLTY019	€67.41	Yes	75	23,919
	Telecommunications	FCLTY008	€70.67	No	150	1,425
		FCLTY025	€68.24	No	50	927
Bruges	Healthcare	FCLTY034	€126.05	No	40	7,233
Brussels	Healthcare	FCLTY035	€227.86	Yes	120	11,706
Copenhagen	Automotive	FCLTY012	€22.03	Yes	75	1,461
	Banking	FCLTY037	€13.69	No	50	0
	Telecommunications	FCLTY002	€34.45	No	40	5,861
		FCLTY037	€13.69	No	50	0
Düsseldorf	Banking	FCLTY029	€2.69	No	50	0
	Construction	FCLTY005	€38.83	No	50	10,623
		FCLTY018	€24.12	No	80	1,907
	Telecommunications	FCLTY029	€2.69	No	50	0
Frankfurt	Construction	FCLTY042	€60.25	No	80	3,872
Gothenburg	Construction	FCLTY032	€81.70	No	150	5,933
	Telecommunications	FCLTY011	€11.34	Yes	60	9,184
Helsinki	Null	FCLTY040	€5.55	No	75	0
	Energy	FCLTY040	€5.55	No	75	5,664
	Technology	FCLTY001	€12.14	Yes	50	0
		FCLTY021	€8.05	Yes	100	8,571

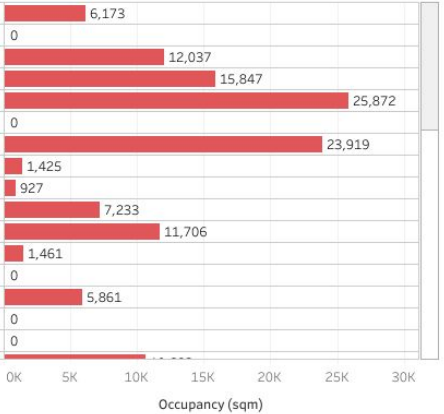
10K 20K 30K 40K

Occupancy (sqm)

Visualization: An Interactive Dashboard

Details

City	Industry	Facility ID	Avg. lease rate/sqm..	BREEAM	Employees.
Aalborg	Construction	FCLTY010	€10.35	No	40
Aarhus	Technology	FCLTY015	€10.28	Yes	60
Amsterdam	Construction	FCLTY004	€8.93	Yes	75
		FCLTY036	€4.71	Yes	75
	Healthcare	FCLTY024	€204.98	Yes	150
Antwerp	Null	FCLTY026	€3.74	Yes	80
Berlin	Energy	FCLTY019	€67.41	Yes	75
	Telecommunications	FCLTY008	€70.67	No	150
		FCLTY025	€68.24	No	50
Bruges	Healthcare	FCLTY034	€126.05	No	40
Brussels	Healthcare	FCLTY035	€227.86	Yes	120
Copenhagen	Automotive	FCLTY012	€22.03	Yes	75
	Banking	FCLTY037	€13.69	No	50
	Telecommunications	FCLTY002	€34.45	No	40
		FCLTY037	€13.69	No	50
Düsseldorf	Banking	FCLTY029	€2.69	No	50



Limitations & Assumptions of Analysis

Limitations

- Merge data by common columns. It would be better to have a ‘CostumerID’ on the Contracts and Leases sheets.
- Does not consider other locations within a certain radius near existing warehouses.
- Does not consider other industries than the industry of previous clients.

Assumptions:

- The lease rate is annual.
- Existing leases have started (as the start date is unknown)
- Exclude the expired leases & contracts in calculating the utilization rate (for accurate facility availability)