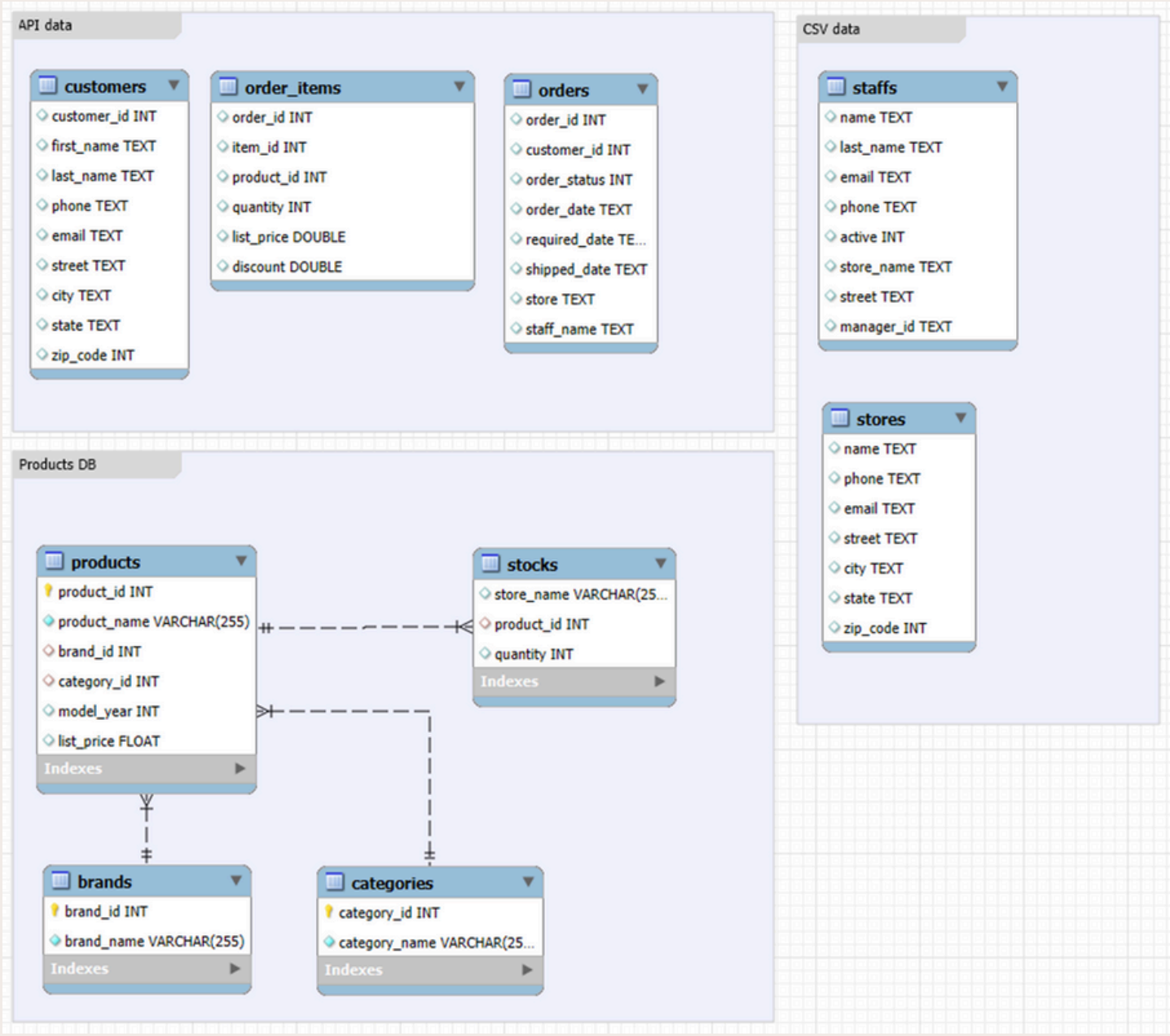

BikeCorp

Tidligere datastruktur



Behov hos BikeCorp

Kunder og geografi

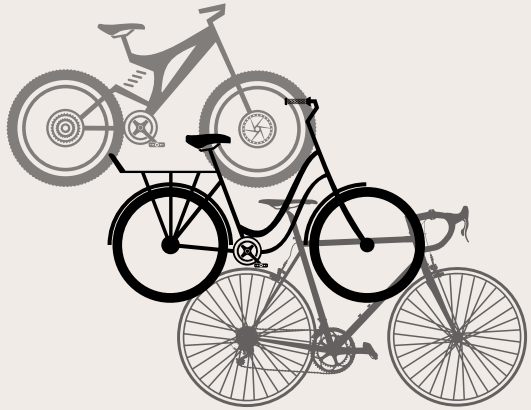
- Kunders lokation sammenholdt med produkter
- Hvilke produkter klarer sig bedst på bestemte markeder

Ordrer og butik

- Årshjul for ordrer; hvilke produkter bliver købt hvornår
- Tidsstempler; hvor lang tid går der fra order date til shipping date
- Hvilke butikker sælger mest af hvilke produkter

Produkter

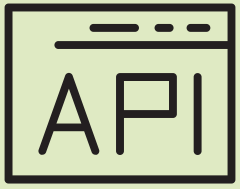
- Hvilke brands klarer sig bedst
- Sammenligning af produktkategorier år til år
- Salgstal



ETL proces

Data før transformation

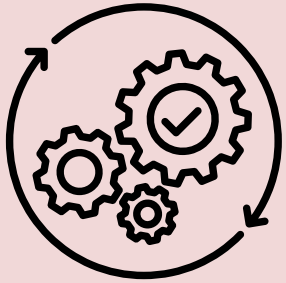
ProductDB



Server/API

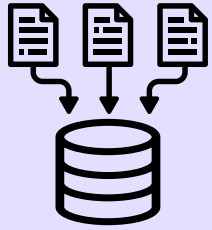
CSV filer

Transformation



- Dataformat ændret og omdøbt for at skabe sammenhæng
- Datoformater er blevet ensrettet
- Data er overført i de samme tabeller for at skabe transparens

Ny database



new_bikecorp_db

- Relationer mellem data er oprettet i MySQL Workbench
- Data i enslydende kolonner har samme navn, så det er muligt at lave joins og merge tabeller

Eksempel på kode

load_orders_from_api.py ● | load_customers_from_api.py ● | load_order_items_table.py

load_orders_from_api.py > ...

```
40
41  import requests
42  import pandas as pd
43  import mysql.connector
44  import json
45
46  response = requests.get("http://127.0.0.1:8000/orders")
47  orders_data = json.loads(response.json())
48
49  orders_df = pd.DataFrame(orders_data)
50
51  if 'store' in orders_df.columns:
52      orders_df = orders_df.rename(columns={"store": "store_name"})
53
54  conn = mysql.connector.connect(
55      host="localhost",
56      user="root",
57      password="Velkommen25",
58      database="new_bikecorp_db"
59  )
60  cursor = conn.cursor()
61
62  for _, order in orders_df.iterrows():
63      cursor.execute("""
64          INSERT INTO orders (
65              order_id, customer_id, order_date, shipped_date, store_name
66          ) VALUES (%s, %s, %s, %s, %s)
67      """, (
68          int(order["order_id"]),
69          int(order["customer_id"]),
70          order["order_date"],
71          order["shipped_date"],
72          order["store_name"]
73      ))
74
75  conn.commit()
76  print(f"{len(orders_df)} ordrer indsat.")
77  cursor.close()
78  conn.close()
79
```

Eksempel på kode

Query 1 | productdb*

Limit to 10000 rows

```
1  USE new_bikecorp_db;
2
3  ALTER TABLE orders
4  ADD CONSTRAINT fk_orders_customers
5  FOREIGN KEY (customer_id) REFERENCES customers(customer_id);
6
7  ALTER TABLE order_items
8  ADD CONSTRAINT fk_order_items_orders
9  FOREIGN KEY (order_id) REFERENCES orders(order_id);
10
11 ALTER TABLE order_items
12 ADD CONSTRAINT fk_order_items_products
13 FOREIGN KEY (product_id) REFERENCES products(product_id);
14
15 ALTER TABLE products
16 ADD CONSTRAINT fk_products_brands
17 FOREIGN KEY (brand_id) REFERENCES brands(brand_id);
18
19 ALTER TABLE products
20 ADD CONSTRAINT fk_products_categories
21 FOREIGN KEY (category_id) REFERENCES categories(category_id);
22
23 ALTER TABLE stocks
24 ADD CONSTRAINT fk_stocks_products
25 FOREIGN KEY (product_id) REFERENCES products(product_id);
26
```

Ny datastruktur