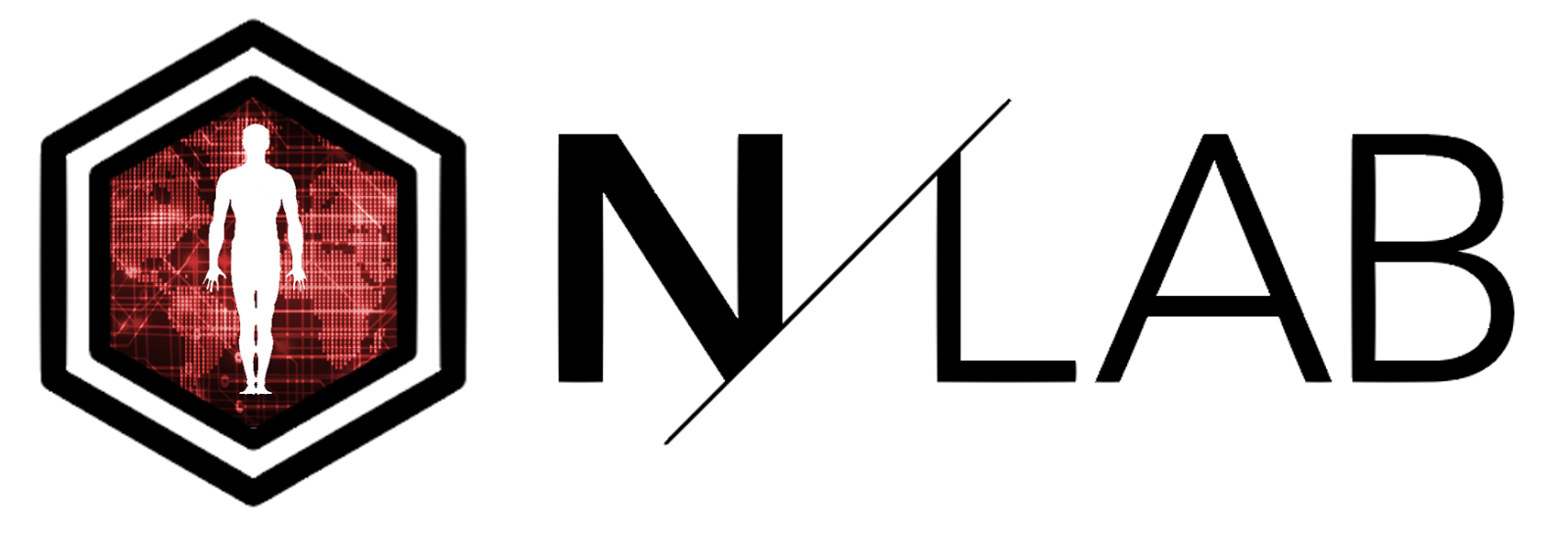
*Data at Scale:* **Coursework**

****

Student ID: 20549904

|  |
| --- |
| **Section 1: The KPIs** |

|  |
| --- |
| **KPI Description:** Sales per Store |
| **KPI formula:** sum(value), per store |
| **Steps to realize KPI:**  1) select store\_code, address, round(sum(value),3) as total\_sales  from receipts  join receipt\_lines  using (receipt\_id)  join stores  using (store\_code)  group by 1,2  order by 3 desc  2) Visualized via Tableau as a graph titled "Sales per store". See the Tableau file.  3) Data underpins Figure 6 in the comparative analysis. |
| **Additional Notes:** None. |

|  |
| --- |
| **KPI Description:** Sales per year per Store |
| **KPI formula:** sum(value), per year, per store |
| **Steps to realize KPI:**  1) select store\_code, date\_format(date\_trunc('year',purchased\_at),'yyyy') as sales\_year, round(sum(value),2) as total\_sales  from receipts  join receipt\_lines  using (receipt\_id)  group by 1,2  order by 1,2  2) Visualized via Tableau as graph titled "Sales per year per store". See the Tableau file. |
| **Additional Notes:** None. |

|  |
| --- |
| **KPI Description:** Sales per quarter per Store |
| **KPI formula:** sum(value), per quarter, per store |
| **Steps to realize KPI:**  1) select store\_code, date\_trunc('quarter',purchased\_at) as sale\_quarter, round(sum(value),2) as total\_sales  from receipts  join receipt\_lines  using (receipt\_id)  group by 1, 2  order by 1, 2  2) Visualized via Tableau as a graph titled "Sales per quarter". See the Tableau file.  3) Data underpins Figure 1 in the comparative analysis. |
| **Additional Notes:** None. |

|  |
| --- |
| **KPI Description:** Month with max sales per store per year |
| **KPI formula:** max(sum(value)), per month, per store |
| Steps to realize KPI:  1) create or replace view MonthlySales as  select store\_code,        extract(month from purchased\_at) as sales\_month,        extract(year from purchased\_at) as sales\_year,        round(sum(value),2) as total\_sales  from receipts  join receipt\_lines  using (receipt\_id)  group by 1, 2, 3  select store\_code, sales\_month, sales\_year, total\_sales  from MonthlySales a  where total\_sales = (    select max(total\_sales)    from MonthlySales b    where a.store\_code = b.store\_code    and a.sales\_year = b.sales\_year  )  order by 1,2 |
| **Additional Notes:** None. |

|  |
| --- |
| **KPI Description:** Active customers per month per store |
| **KPI formula:** count(unique customers), per month, per store |
| **Steps to realize KPI:**  1) select store\_code, date\_format(date\_trunc('month',purchased\_at),'MM-yyyy') as Month\_year, count(distinct customer\_id) as active\_customers  from receipts  group by 1, 2  order by 1, 2  2) Visualized via Tableau as graph titled "Active Customers per month per store". See the Tableau file.  3) Data underpins Figure 4 in the comparative analysis. |
| **Additional Notes:** None. |

|  |
| --- |
| **KPI Description:** First-time customers per month per store |
| **KPI formula:** count(unique customers), min(purchased\_at), per month, per store |
| **Steps to realize KPI:**  1) select store\_code,  date\_format(date\_trunc('month', first\_purchase),'MM') as first\_purchase\_month,  date\_format(first\_purchase,'yyyy') as first\_purchase\_year,  count(distinct customer\_id) as new\_customers\_count  from (  select distinct store\_code, customer\_id, min(purchased\_at) as first\_purchase  from receipts  group by 1,2)  group by 1,2,3  order by 1,2  2) Visualized via Tableau as a graph titled "New customers per month per store". See the Tableau file.  3) Data underpins Figure 5 in the comparative analysis. |
| **Additional Notes:** None. |

|  |
| --- |
| **KPI Description:** Repeat customers per month per store |
| **KPI formula:** count(unique customers), per month, per store |
| **Steps to realize KPI:**  1) select store\_code, purchase\_month, count(\*) as total\_repeat\_customers  from (  select distinct customer\_id, store\_code, date\_trunc('month', purchased\_at) as purchase\_month  from receipts  group by 1, 2, 3  having count(distinct purchased\_at::DATE) > 1  order by 1)  group by 1, 2  order by 1, 2 |
| Additional Notes: Customers visiting the store more than once a month on separate days. |

|  |
| --- |
| **KPI Description:** Active repeat customers per store |
| **KPI formula:** count(unique customers), per month, per store |
| Steps to realize KPI:  1) create or replace view RepeatCustomersPerMonthPerStoreTenPurchase as  select distinct customer\_id, store\_code, date\_trunc('month', purchased\_at) as purchase\_month  from receipts  group by 1, 2, 3  having count(distinct purchased\_at::DATE) > 10  order by 2  create or replace view CustomersWithRecentPurchaseIn2022 as  select distinct customer\_id, store\_code, date\_trunc('month',max(purchased\_at)) as recent\_purchase  from receipts  group by 1, 2  having EXTRACT(YEAR FROM MAX(purchased\_at)) = 2022  select r.store\_code, count(distinct r.customer\_id) as active\_repeat\_customers  from RepeatCustomersPerMonthPerStoreTenPurchase r, CustomersWithRecentPurchaseIn2022 c  where r.customer\_id = c.customer\_id  and r.store\_code = c.store\_code  group by 1  order by 1 |
| Additional Notes: Customers who have been visiting the store more than 10 times with recent purchase made in the year 2022. |

|  |
| --- |
| **KPI Description:** Total High-value customers per store |
| **KPI formula:** customer spend > avg spend for store , per store |
| **Steps to realize KPI:**  1) create or replace view avg\_sales\_per\_store as  select store\_code, sum(value) as total\_sales,        count(distinct customer\_id) as total\_customers,        sum(value)/count(distinct customer\_id) as avg\_sales  from receipts  join receipt\_lines  using (receipt\_id)  group by 1  create or replace view total\_customer\_sales as  select distinct customer\_id, store\_code, sum(value) as total\_customer\_sales  from receipts  join receipt\_lines  using (receipt\_id)  group by 1,2  select store\_code, count(distinct customer\_id) as total\_high\_value\_customers  from total\_customer\_sales  join avg\_sales\_per\_store  using (store\_code)  where total\_customer\_sales >= avg\_sales  group by 1  order by 1 |
| Additional Notes: Customers who spend at least average spend value for each store. |

|  |
| --- |
| **KPI Description:** Total Low-value customers per store |
| **KPI formula:** customer spend < avg store spend, per store |
| **Steps to realize KPI:**  1) select store\_code, count(distinct customer\_id) as total\_low\_value\_customers  from total\_customer\_sales  join avg\_sales\_per\_store  using (store\_code)  where total\_customer\_sales < avg\_sales  group by 1  order by 1 |
| Additional Notes: Customers who spend less than the average spend value for each store. |

|  |
| --- |
| **KPI Description:** Average transaction value of customers per quarter per store |
| **KPI formula:** avg customer spend, per quarter, per store |
| **Steps to realize KPI:**  1) select store\_code, date\_trunc('quarter', purchased\_at) as purchase\_quarter,        round(sum(value),2) as total\_spend\_per\_quarter,        count(distinct customer\_id) as customers\_per\_quarter,        round(sum(value)/count(distinct customer\_id),2) as avg\_spend\_per\_quarter  from receipts  join receipt\_lines  using (receipt\_id)  group by 1,2  order by 1,2  2) Visualized via Tableau as a graph titled "Average Transaction value per store per quarter". See the Tableau file.  3) Data underpins Figure 7 in the dashboard for comparative analysis. |
| Additional Notes: Customers visiting the store more than once a month on separate days. |

|  |
| --- |
| **KPI Description:** Customer retention |
| **KPI formula:** avg(months), per store |
| **Steps to realize KPI:**  1) create or replace view monthsCustomerStaysPerStore as  select distinct customer\_id,        store\_code,        min(purchased\_at) as first\_purchase\_date,        max(purchased\_at) as most\_recent\_purchas\_date,        months\_between(max(purchased\_at), min(purchased\_at)) as total\_months\_as\_customer  from receipts  group by 1,2  order by 1  select store\_code,         round(AVG(total\_months\_as\_customer),2) AS avg\_months\_customer\_stays  from monthsCustomerStaysPerStore  group by store\_code  order by 2 desc  2) Visualized via Tableau as a graph titled "Average months for customer retention per store". See the Tableau file.  3) Data underpins Figure 3 in the comparative analysis. |
| Additional Notes: Average no. of months the customer continues buying in a store i.e continues to be a customer of that store. |

|  |
| --- |
| **KPI Description:** Average customer lifetime value per store |
| **KPI formula:** avg(purchase\_value) \* avg(purchase\_frequency) \* avg(customer\_lifespan\_months), per store |
| **Steps to realize KPI:**  1) select store\_code, avg(clv) as average\_customer\_life\_time\_value\_per\_store  from (  select distinct customer\_id, store\_code,         avg\_purchase\_value,         avg(purchase\_frequency) as average\_purchase\_frequency,         avg(customer\_lifespan\_months) as average\_customer\_lifespan,         avg\_purchase\_value \* avg(purchase\_frequency) \* avg(customer\_lifespan\_months) as clv  from (      select customer\_id, store\_code,            sum(value)/count(distinct receipt\_id) as avg\_purchase\_value,            count(distinct receipt\_id) as purchase\_frequency,            months\_between(max(purchased\_at), min(purchased\_at)) as customer\_lifespan\_months      from customers            join receipts            using (customer\_id)            join receipt\_lines            using (receipt\_id)      group by 1, 2)  group by 1,2,3)  group by 1 |
| Additional Notes: None. |

|  |
| --- |
| **KPI Description:** Best Selling Products per Store by revenue |
| **KPI formula:** product\_code, max(sales), per store |
| **Steps to realize KPI:**  **1)** create or replace view RevenueAndQuantityOfProductPerStore as  select store\_code, product\_code, product\_details, category\_details,          round(sum(value),2) as product\_sales, sum(qty) as product\_unit\_count  from receipt\_lines  join products  using (product\_code)  join receipts  using (receipt\_id)  group by 1,2,3,4  order by 5 desc  select store\_code, product\_code, product\_details, product\_sales,product\_sales/product\_unit\_count as price\_of\_product  from RevenueAndQuantityOfProductPerStore a  where product\_sales = (    select max(product\_sales)    from RevenueAndQuantityOfProductPerStore b    where a.store\_code = b.store\_code  ) |
| Additional Notes: Average no. of months the customer continues buying in a store i.e continues to be a customer of that store. |

|  |
| --- |
| **KPI Description:** Best Selling Products per Store by quantity |
| **KPI formula:** product\_code, max(value), per store |
| **Steps to realize KPI:**  **1)** select store\_code, product\_code,product\_details, category\_details, product\_unit\_count,         round(product\_sales/product\_unit\_count,2) as price\_of\_product  from RevenueAndQuantityOfProductPerStore a  where product\_unit\_count = (    select max(product\_unit\_count)    from RevenueAndQuantityOfProductPerStore b    where a.store\_code = b.store\_code  )  order by 1 |
| Additional Notes: View created during calculating best-sellers by revenue is used here. |

|  |
| --- |
| **KPI Description:** Most popular department per store |
| **KPI formula:** department\_code, max(value), per store |
| **Steps to realize KPI:**  1) create or replace view departmentSalesPerStore as  select department\_code, department\_name, store\_code, sum(value) AS total\_sales  from products  join receipt\_lines  using (product\_code)  join receipts  using (receipt\_id)  group by 1, 2, 3  order by 4 desc  select store\_code, department\_code, department\_name, round(total\_sales,2) as total\_sales  from departmentSalesPerStore a  where total\_sales = (    select max(total\_sales)    from departmentSalesPerStore b    where a.store\_code = b.store\_code  )  order by 1 |
| Additional Notes: None. |

|  |
| --- |
| **KPI Description:** Most popular category per store |
| **KPI formula:** category\_code, max(value), per store |
| Steps to realize KPI:  1) create or replace view categorySalesPerStore as  select category\_code, category\_details, store\_code, department\_name,         sum(value) AS total\_sales\_category  from products  join receipt\_lines  using (product\_code)  join receipts  using (receipt\_id)  group by 1, 2, 3, 4  order by 5 desc  select store\_code, category\_code, category\_details, department\_name,        round(total\_sales\_category,2) as total\_sales\_category  from categorySalesPerStore a  where total\_sales\_category = (    select max(total\_sales\_category)    from categorySalesPerStore b    where a.store\_code = b.store\_code  )  order by 1 |
| Additional Notes: None. |

|  |
| --- |
| **Section 2: Comparative Analysis**  (including recommendations) |

FoodCorp is a medium-sized dynamic company with a presence in Nottingham, Birmingham, and London. Being in operation for less than 2 years, they plan to strategically focus marketing efforts on one of their stores. Analyzing customer behavior across different store locations will help tailor the marketing campaign effectively, increasing the likelihood of success.

**Customer Metrics Overview:**

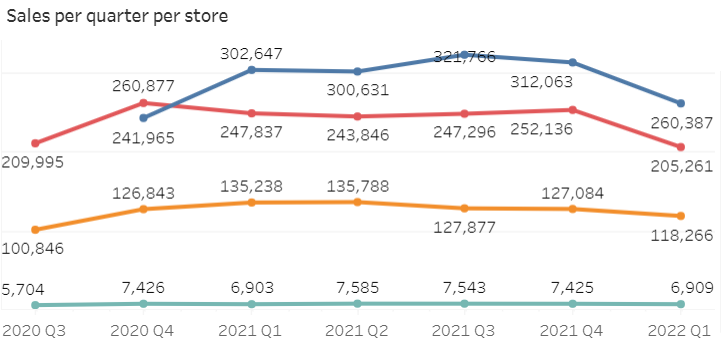
FoodCorp has a strong presence with a customer base of 11,474 and a diverse range of products comprising 21,812 unique items. However, the initial analysis reveals that 156 products have never been sold in any stores. This highlights the need to develop meticulous strategies to introduce these products to the market.

**Data Integrity and Missing Values:**

A total of 9907 records of customers lack ‘Date\_of\_birth’ values rendering it impractical to determine customer trends by age as a substantial amount of data is missing. While the products table maintains data integrity, certain values of department\_code provide information about unknown categories and sub-categories. They are mentioned with department\_code = -2. In receipt\_lines, 3004 product\_codes and 148 ‘qty’ fields are null, with 657 instances of negative values in the ‘value’ and ‘qty’ fields signaling returned products and associated refunds.

**Store Characteristics and Size:**

Adequate data to calculate the size of the store is not available, thus, an assumption is made considering that the store with the maximum no. of products will have more space and will be the largest. On analysis, it is perceived that Store 0 is the largest (127,366 products by quantity), while Store 3 is the smallest (44,152 products).



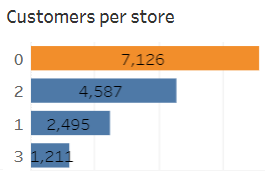


Fig 1Fig 2

**Customer Behaviour and Retention:**

Store 0 has cultivated a strong customer base of 7,126 customers, leading the way, followed by Store 2 and Store 3 with 4,587 and 2,495 customers respectively (refer to Fig 2). Notably, Store 3 exhibits the lowest number with 1,211 overall customers.

Store 0 boasts the highest number of 3,346 active customers in the year 2022 with the average months for customer retention being 7.71 months. In contrast, Store 3 documents the lowest customer count at 343, with the shortest customer retention at 5.51 months. Store 2 and Store 1 have 2,055 and 1,116 active customers in 2022, respectively, with Store 2 exhibiting the highest overall customer retention, averaging 9.21 months (refer to Fig 3).

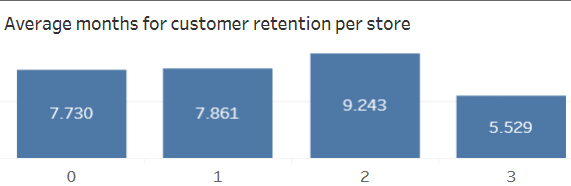


Fig 3

A similar trend in new customer influx per month is observed, where despite starting operations in October 2020, two months later than the rest, store 0 has recorded the highest number of customer joinings, while Store 3 has the lowest (refer to Fig 5).

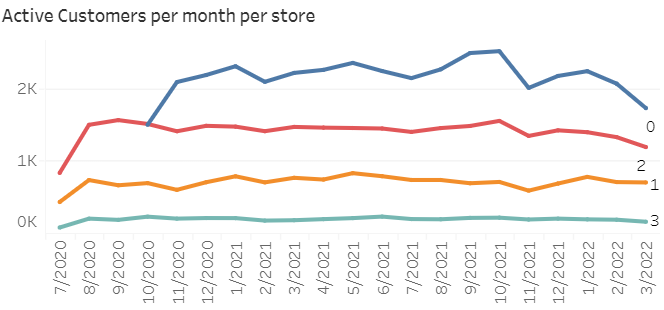
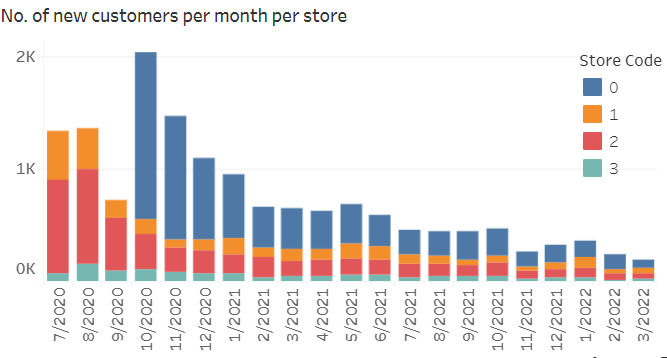


Fig 4 Fig 5

When examining the average transaction values for each store across quarters, a consistent trend emerges, with Store 1 and Store 2 consistently maintaining values above and around £100(refer to Fig 7). In contrast, Store 3 consistently reports the lowest values in this regard. Notably, the average customer lifetime value for Store 1 stands as the highest at £6,116.57, closely followed by Store 2 at £6,090.98. However, Store 0 and Store 3 fall behind in this aspect. The store with the most high-value customers is Store 0, boasting a total of 1,685 customers, with Store 2 following closely behind with 1,119 customers.

**Sales Performance:**

Store 0 leads with £1,729,458.95 overall sales(refer to Fig. 6), while Store 3 records the lowest £ 49,495.21. Analyzing the sales pattern per month reveals that October has consistently brought the highest monthly sales for all stores.

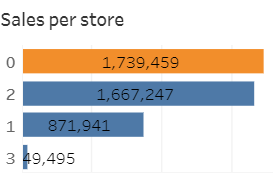
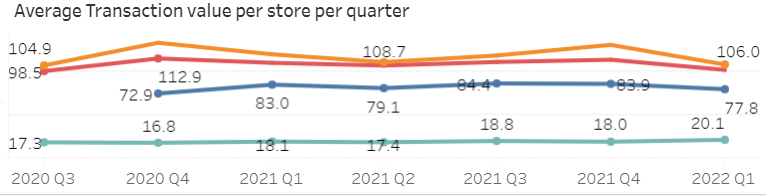


Fig 6 Fig 7

**Product Insights:**

Top-selling products of FoodCorp across all stores are- 'Pay Point Each', 'Fuel Sales Each', 'National Lotto', 'PayPoint Powerc', and 'Home Brand 4Pint fres'. The product with code 9357 (Pay Point Each) emerges as a best-seller across all stores in terms of sales revenue. However, pricing discrepancies are notable here, particularly in London stores.

The best sellers by quantity varies, with 'Bananas' in Fruits category leading in Store 0, Store 1, and Store 2, while Store 3 sees the highest quantity of milk sales. Notably, all these best-sellers by quantity have an average pricing of 0.67 pounds.

Thus, it explains varied customer behavior influenced by product categories and sales.

**Department wise analysis:**

Store 0 and Store 2, with 42 departments exhibit a broader range of products being sold compared to Store 1 (39 departments) and Store 0 (33 departments). Popular departments vary from ‘Provisions Dairy’ in Store 0, ‘Cigarettes and Tobacco’ focused in Store 1 while London stores prioritizing ‘Fruits and Vegetables’.

Upon a comprehensive evaluation of all stores, it is evident that Store 0 in Nottingham has exhibited superior performance, whereas Store 3 has performed less favorably in terms of sales and various customer metrics. However, when scrutinizing crucial factors such as average customer lifetime value and average transaction values per store, pivotal for establishing robust customer retention, Store 0 falls short. Upon assessing analogous metrics for Store 2 at 14 Handyside St in London, it becomes apparent that it holds greater potential for strategic growth. Consequently, I recommend Store 2 for the upcoming marketing campaign. The supporting arguments for the same are as:

1. **Market Potential:**

London, being bustling and populous city, offers an expansive market. Store 2 stands as 2rd largest store with 18,363 distinct products available, however, the customers here are just 4,785. Executing a marketing campaign here provides a good opportunity for FoodCorp to attract potentially untapped customers for new customer acquisition and enhanced market presence.

1. **Customer Behaviours:**

* Customers at Store 2 demonstrate a significant average transaction value, amounting to £363.47, indicating a propensity for substantial spending.
* The store boasts 1,119 high-value customers, which constitute a significant one-fourth of its total customer base.
* Moreover, these customers exhibit an exceptional level of loyalty, evidenced by the highest customer retention period of 9.21 months.
* The average customer lifetime value at Store 2 ranks as the second-highest, reaching £6,090.99. These robust customer attributes position London, particularly Store 2, as a highly suitable target for initiatives aimed at acquiring new customers and boosting sales.

1. **Upselling Opportunities:**

* This Store has an unexplored customer base with a high range of diverse products available in the store leading to significant upselling opportunities. Existing customers can be encouraged to explore new products and try for higher-value products through new tailored offers.
* The popularity of the 'Fruits and Vegetables' department suggests a health-conscious consumer base. This aligns seamlessly with the prevailing global emphasis on adopting a healthy lifestyle, thereby serving as a pivotal aspect for promoting the upselling of new products with enhanced value within this category.

Considering all the above points, I believe store 2 has the highest potential to explore and improve FoodCorp’s presence in the UK market.