BI Support Engineer Applicant Tasks

Download Excel file: [data\_extract\_sales\_201612.xlsx](https://docs.google.com/spreadsheets/d/13jrXmXqi7giOsUYspRIzI-MHudGiJNj_/edit?usp=share_link&ouid=103165194570893739539&rtpof=true&sd=true)

The file contains a sample of 2000 line items from an imaginary grocery seller “Frozen Fresh, Ltd.”

The columns of interest:

* PRODUCT\_CATEGORY\_NAME
* TRANSACTION\_NUMBER
* QUANTITY

You will need to calculate and report multiple KPIs out of it.

Average basket is a metric that describes common traits in customer behavior such as how much they spend, how many items they buy &c.

**1. Calculate how many items on average per transaction was purchased within each category? (2x turkey breast + 1x jumbo wings = 3 items).**

**2. Calculate how many unique items on average per transaction was purchased within each category (2x turkey breast + 1x jumbo wings = 2 items).**

**3. Calculate the number of unique receipts (number of transactions) within each category. The receipt may contain products from different categories.** In this case the receipt may be split (or shared) between categories:

\* If a receipt contains 1x turkey breast and 1x shrimp cocktail it counts as two receipts: 1 for poultry and 1 for seafood.

\* If a receipt contains 1x turkey breast and 1x jumbo wings it counts as one receipt: poultry

Sometimes we want to research the buying patterns for various reasons. Consider a scenario where most people will buy 1 or 2 items at a time, occasionally 3 or 4. These purchases typically make up 95% of the trade volume. Note that these numbers are just illustrative, you will need to calculate the actual number that fits the 95% pattern. It may be 1-2, 1-3, 1-4 or something else.

We need to know how many transactions are not within the 95% “common” purchase size. One way of doing it is ZScore.

**4. Which ZScore value is correct to filter the 5% of exceptions?**

**5. What is the actual number of 5% exceptions within the data sample?**

**6. Produce a nicely formatted, clear and concise report for each of the 1-3 questions presentable on a daily basis to the management team of the above company. For 4 and 5 the answer is one-liner.**

You can employ different techniques: Excel, Python, R, or any other tool you feel comfortable with.

**7. Write a few lines about the technique you were using, any problems you encountered. If you were unable to find an answer, write also the reasons why you think you could not, or state so in the respective report (which will have no values in this case).**

If you have any questions regarding the tasks please ask [**applicantfeedback@targetta.net**](mailto:applicantfeedback@targetta.net)during the time you have been assigned.

Submit your work here: <https://forms.gle/overFR4CjtqgxULk6>