**Strategy Document: Google Fiber**

**Sign-off matrix:**

| **Name** | **Team / Role** | **Date** |
| --- | --- | --- |
|  |  |  |

**Proposer:** Hiring Manager

**Status:** Implemented

**Primary dataset:** contains the needed metrics to respond to the stakeholders questions, including:

* The type of internet service received (Market\_city) : no\_internetservice, DSL, Fiber\_optic
* Contract : One year, Two year, Month-to-month
* TechSupport : No internet service, Yes, No
* The average repeat calls related to combined problem types (total\_problem\_type): integer
* The average repeat calls related to the problem type of account management (account\_mang\_problem\_type): integer
* The average repeat calls related to the problem type of technician troubleshooting (tech\_troubleshoot\_problem\_type): integer
* The average churn or customer retention: boolean (True=customers stop service, False=customers keep service)
* The average frequency of repeat calls related to combined problem types [Contract/number of tickets] (total\_problem\_type\_call\_freq): integer
* The average frequency of repeat calls related to the problem type of account management (accmang\_problem\_type\_call\_freq): integer
* The average frequency of repeat calls related to the problem type of technician troubleshooting (techtrob\_problem\_type\_call\_freq): integer

A BigQuery SQL database was constructed from a single Kaggle .csv file. Automatic ingestion was performed via the script <https://github.com/j622amilah/Case_Studies/tree/main/3_case_study_Google_fiber/case_study_main.sh> and self-written GCP library <https://github.com/j622amilah/GCP_ingestion_analysis_tools/blob/main/GCP_bigquery_case_study_library.sh> . Two basic queries were performed to create a final aggregated SQL database with the listed columns above for visualizing the results. A Google sheet was created and Data - Data Connectors - Connect to BigQuery were selected to connect to the BigQuery database.

**Secondary dataset:** none

## User Profiles :

## The intended audience for this dashboard is a technical audience including the Hiring Manager, Project Manager, Lead BI Analyst, and BI Analysts. The intention for the dashboard is to give answers for the three asked questions:

1. How often does the customer service team receive repeat calls from customers?
2. What problem types generate the most repeat calls?
3. Which market city’s customer service team receives the most repeat calls?

# **Dashboard Functionality**

A dashboard was constructed from a BigQuery Google connected sheet using Google Apps Script GET Request doGet function; the web app is located at <https://script.google.com/macros/s/AKfycbzdatlcxME8jF3jpwmzfizMkn1Z-9DDBhjOXM9aknaZbAdWnpygUYAmQgX7yA31KfUa/exec> for anyone to view updated information regarding the BigQuery SQL database. This dashboard solution is automatic and directly connected to the BigQuery SQL database; the diagrams change when the BigQuery SQL database has updated changes.

A dashboard could have been created using several methods: Looker/LookML, Tableau, Google Apps Script, Cloud Function. Connecting the BigQuery database to a Google connected sheet and deploying the connected sheet with Google Apps Scripts was faster than the procedure to connect the BigQuery database to Looker/LookML, Tableau, and Cloud Function. The Looker/LookML, Tableau, and Cloud Function methods required more time investment to learn how to use their specific protocol and/or language. Google Apps Scripts uses a JavaScript derived language called Google Scripts and HTML thus coding the deployment functionality was with well-known familiar tools.



# **Metrics and Charts**

Please create a table like the example below for each chart that you’d like to include in the dashboard. If you’d like to break the dashboard under different headers, feel free to list those here as well.

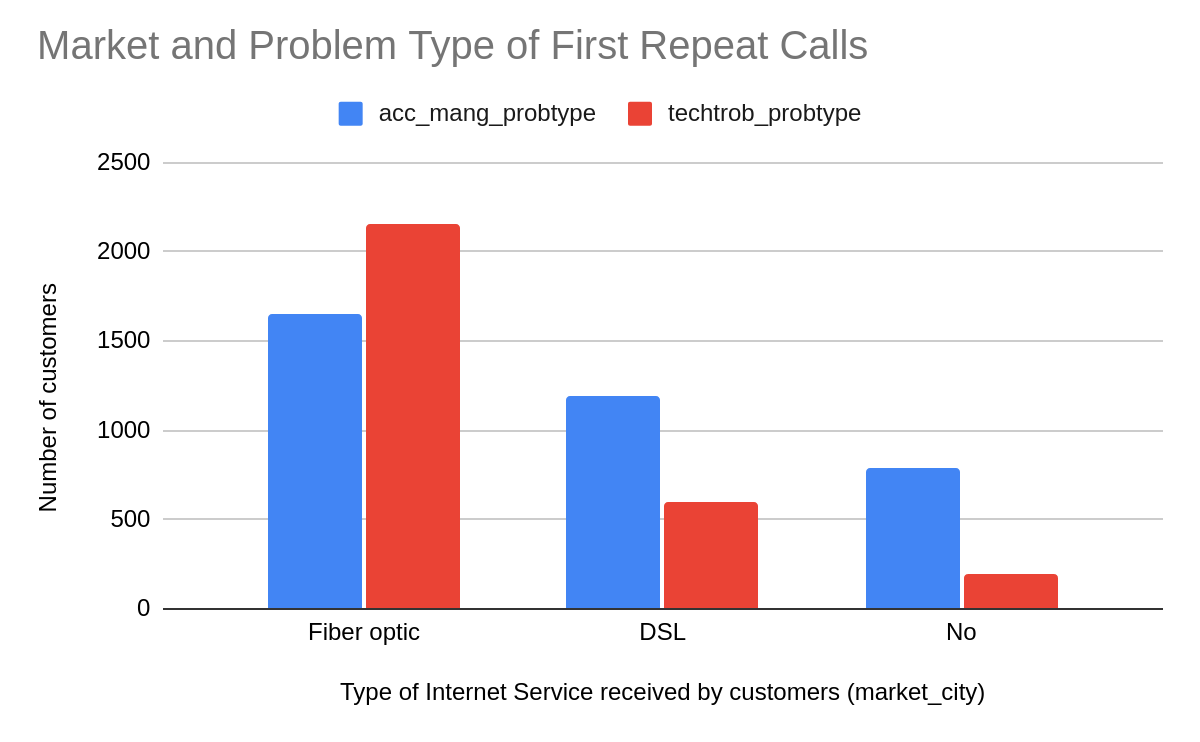
### Chart 1

In the Kaggle dataset, there is no information about which date a ticket/call was received, it only indicates the number of problem type tickets made by each CutomerId. Unfortunately, one can not create a chart of calls per weekday.

### Chart 2

This chart shows the number of customer repeat calls per Internet Service type, called market\_city. Customer calls were included in the query if the number of calls made was greater than 1, indicating that they made multiple or “repeat” calls with respect to their Internet Service subscription. Repeat calls per issue was not evaluated because the given database did not indicate how many calls were linked to the same issue, it only gave the number of calls/tickets with respect to a problem type. There was no date information given per call, thus I could not calculate whether calls were successive, indicating that perhaps customers called about the same issue.

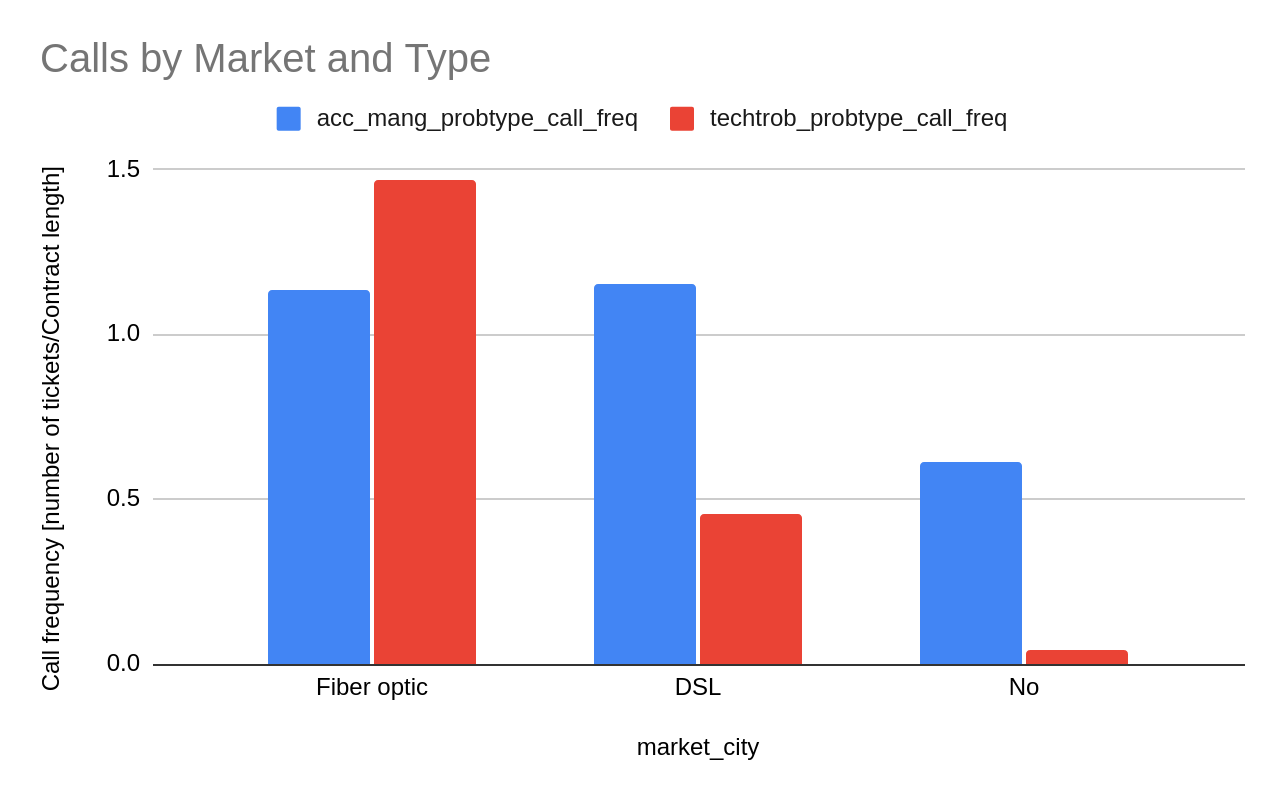
The results of evaluating repeating calls with respect to Internet Service subscription show that Fiber optic customers call the most about technician troubleshooting problems. Over 2000 customers of the 7000+ customers repeatedly called about technical issues.



### Chart 3

This chart shows the frequency of calling per Internet Service type (market\_city). Due to the fact that the dataset did not have DateTime information per call/ticket, a frequency measure was calculated using the number of calls per Contract length for each problem type.

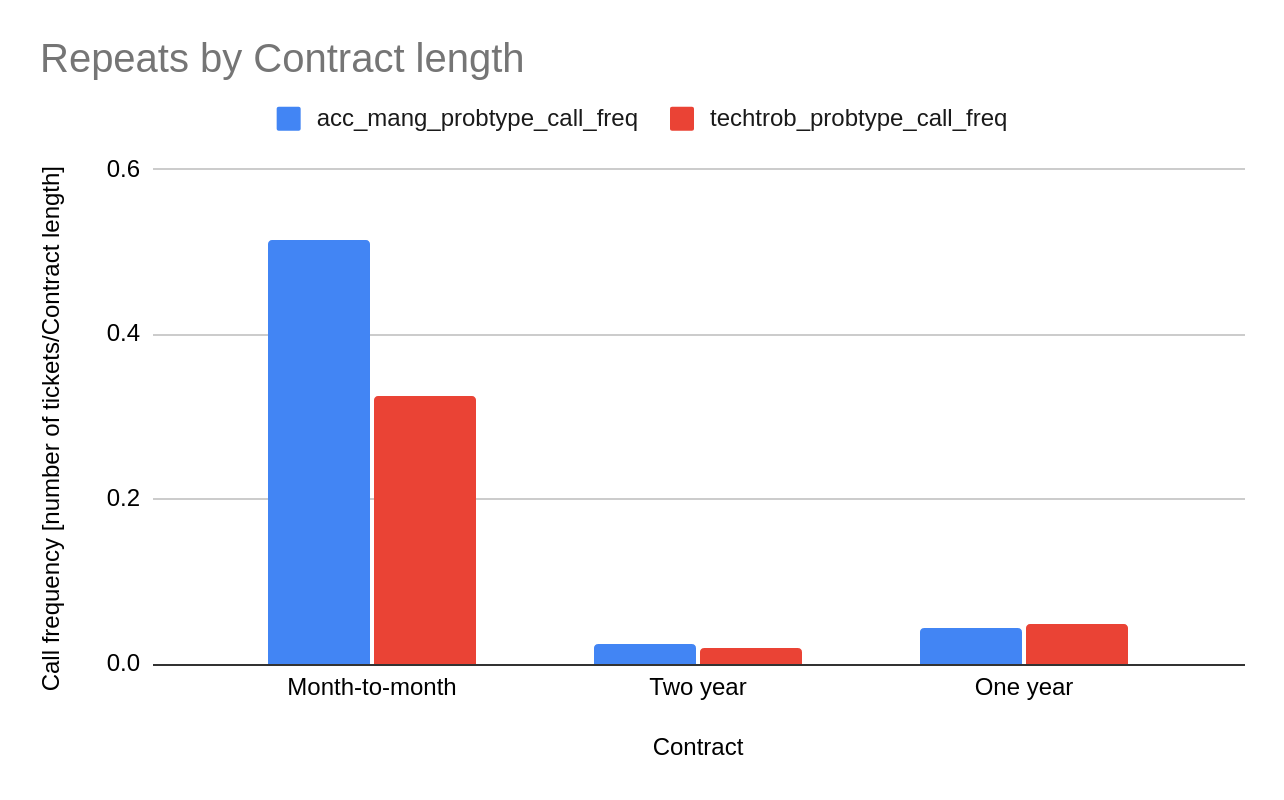
The results of evaluating the frequency of repeating calls with respect to Internet Service subscription show that Fiber optic customers call most frequently about technician troubleshooting problems, whereas DSL and non-Internet Service customers call most frequently about account management issues.



### Chart 4

This chart shows the frequency of calling per customer Contract type. Again, due to the fact that the dataset did not have DateTime information per call/ticket, a frequency measure was calculated using the number of calls per Contract length for each problem type.

The results of evaluating the frequency of repeating calls with respect to Contract type show that Month-to-month customers call most frequently about account management and technician troubleshooting issues in comparison to one-year and two-year Contract customers.



In summary, considering the three dashboard charts, Month-to-month Fiber Optic customers repeatedly and frequently call the most out of any of the population groups. In addition, these Month-to-month Fiber Optic customers are the only customers to quit their service, indicated by an average churn greater than 0.5 indicating True. The second most repeated calling population type was non-Fiber optic customers for account management issues, then Fiber optic customers for technician troubleshooting.

