**My Answers**

**Part 1.**

(CMPM \* MEMBER COUNT) \* (MRR / (1+IR-MRR))

Please note that as someone who has not worked in the healthcare industry, I cannot guarantee the accuracy of this formula. However, based on my research and understanding, it appears that the formula (CMPM \* MEMBER COUNT) \* (MRR / (1+IR-MRR)) can be used to estimate the lifetime value of a particular cohort, taking into account factors such as commission per member, retention rates, interest rates, and the number of members in the cohort.

The first part, (CMPM \* MEMBER COUNT), calculates the total revenue earned from the cohort. It multiplies the commission earned per member each month by the number of members in the cohort to get the total revenue earned over the cohort lifetime.

The second part, (MRR / (1+IR-MRR)), calculates the discount rate-adjusted retention rate for the cohort. It takes the monthly retention rate for the cohort and adjusts it for the time value of money using the interest rate. This calculates the probability that a customer will continue to use the service in the future, given the discount rate and retention rate.

Multiplying these two parts together gives you the estimated total value of the cohort over its lifetime.

**Part 2.**

**2-1.**

After analyzing the data, I have found three key conditions that impact the average member lifetime.

Firstly, depending on the state in which the members live, the average cohort lifetime can either be positively or negatively impacted. Idaho state has shown to have a greater average cohort lifetime than other states, and rural or suburb areas tend to have superior policyholders.

Secondly, the carrier that writes a policy contract also has a significant impact on the average member lifetime. In particular, Carrier 11 has outperformed other carriers in terms of the average member lifetime.

Lastly, seasonality plays a crucial role in the average member lifetime. Between Q1 and Q2, both the retention rate and average member lifetime appear to be positive. It is important to note that seasonality may vary depending on the industry and other factors, but it is worth considering when analyzing the average member lifetime.

To support my statements above, I also made a linear regression model (dependent variable Y = MEMBER LIFETIME / independent variables X = Members, Max\_duration, start\_month, start\_year.

Per regression model, the outcome was :

Intercept: -20596.095191480483

MEMBERS: -3.1301856859989887

MAX\_DURATION: 0.24989741764770923

START\_MONTH: -0.3046851503414685

START\_YEAR: 10.236469283347638

The intercept value of -20596.09 means that the predicted value of Y (MEMBER LIFETIME) is expected to be around that value when all the independent variables (MEMBERS, MAX\_DURATION, START\_MONTH, START\_YEAR) are set to 0.

For each additional member, the predicted value of Y is expected to decrease by 3.13.

For each additional unit of MAX\_DURATION, the predicted value of Y is expected to increase by 0.25.

For each additional month in START\_MONTH, the predicted value of Y is expected to decrease by 0.30.

For each additional year in START\_YEAR, the predicted value of Y is expected to increase by 10.24.

**2-2.**

Chart, histogram

Description automatically generated

Chart

Description automatically generated

To answer the question about the typical shape of a monthly retention curve, I conducted extensive research and utilized pivot tables to calculate the retention rates for different cohorts. Based on this analysis, I was able to generate three significant pivot tables to find the retention rate. Using these tables, I created two figures, one representing the monthly retention curve and the other showing the cumulative retention rate. The monthly retention curve shows a typical pattern where the retention rate decreases gradually over the policy duration and eventually flattens out at around 8-9 months. However, it's important to note that the shape of the curve can vary depending on the industry and customer behavior.

**2-3.**

Chart

Description automatically generated

To add on to my previous comment on the retention rates, as I analyzed the line plot of the retention rates for the 2019 cohort group, I noticed that the rates follow a comparable pattern to the previous years' cohort groups. Utilizing this information, I can generate a prediction for the cumulative retention curve of the 2019 cohort group by incorporating previous years' data. Nevertheless, it is imperative to note that the accuracy of this prediction against the actual curve cannot be determined without conducting a comparative analysis and calculating the prediction error. Additionally, I am interested in learning more about this topic and would like to explore it further with the professionals. Please let me know if you have specific tips or suggestions for approaching similar problem.