Final Project

James Beall

American Public University System

BUSN661: Advanced Analytics II

Dr. Xiaodong Wu

August 25, 2024

**Introduction**

GreenTech Innovations is a sustainable technology solutions company that is interested in reversing the trend of stagnant growth over the past year. GreenTech Innovations has collected data on sales, expenses, operations, competitors’ market shares, industry growth by year dating back to 2019. Additionally, GreenTech Innovations has provided summary data on customer feedback as well as a probability model for various market scenarios. GreenTech Innovations has tasked us, the data analyst firm, to analyze their data and make data-driven recommendations to consult the company. The objective is to find decisions that will increase profitability in a rapidly changing market with emerging strong competition.

**Analysis and Correlation**

This section will give an overview of the analysis conducted on the available data related to sales, expenses, operations, competitors’ market shares, and industry growth. Correlations between different datasets will be highlighted and explained.

**Analysis**

Revenue was calculated using sales as income. Revenue on average has a slight positive trend from 2019 to 2023 but this trend is hardly steady. With an R-Squared value of 0.1288 the slight positive trend is hardly linear.

Sales data shows that on average total sales are relatively flat from 2019 to 2023. There is a low in 2020 and a high in 2021 and sales steadily decline from then. The R-Squared value is 0.012 reinforcing the fact that there is no strong linear increase or decrease in sales year over year.

More useful information is derived from analysis of a per-product correlation calculation which will follow.

Expenses totals peak in years 2020 and 2021 and then fall off through 2023. A breakdown of expenses into marketing expenses, operational expenses, and R&D expenses clearly show that the most dramatic decline in expenses from 2022 to 2023 are due to a decline in operational expenses.

Competitor market shares as a total do not follow a linear trend. The R-Squared value is 0.0854. Considering the competitors individually, we see that competitor two has increased market shares over time and competitor three has decreased market shares over time. Competitor one, as of 2023, is nearly exactly between competitor two and three as far as total market shares.

Industry growth and GreenTech Innovations market shares have both declined from 2019 to 2023. Industry growth has declined more steadily with an R-Squared value of 0.8789 compared to GreenTech Innovations’ 0.5298. This suggests that GreenTech Innovations’ market share follows industry growth (or decline) but correlation calculations suggest otherwise.

Operational data shows a more interesting trend. Employment varies little from 2019 to 2023 with a maximum difference of 15 employees from 2019 to 2020 and 2020 to 2021. More interestingly, item returns vs queries shows a strong trend; more queries correspond to less returns. The next section explores correlations with a focus on queries.

**Correlation**

Correlation analysis was run on each pair of data points available. This is a small section of the full table:

|  |  |  |  |
| --- | --- | --- | --- |
|  | *Year* | *Product\_A\_Sales* | *Product\_B\_Sales* |
| *Year* | 1 |  |  |
| *Product\_A\_Sales* | 0.296886551 | 1 |  |
| *Product\_B\_Sales* | 0.843157324 | 0.146998636 | 1 |
| *Product\_C\_Sales* | -0.764953018 | 0.041424693 | -0.39402994 |
| *Total Sales* | 0.109406749 | 0.727652481 | 0.360290099 |
| *Marketing\_Expenses* | 0.141046099 | 0.253774468 | 0.162316281 |
| *Operational\_Expenses* | -0.286861379 | 0.562782366 | -0.547636583 |
| *R&D\_Expenses* | -0.603735529 | -0.161513985 | -0.246938352 |
| *Total Expenses* | -0.592820094 | 0.28964082 | -0.508574428 |
| *Competitor1\_Market\_Share* | -0.431388591 | 0.034216702 | -0.164261977 |
| *Competitor2\_Market\_Share* | 0.668921931 | 0.636085948 | 0.731254068 |

Starting with queries, we noted in the previous section that queries had an inverse correlation with product returns. Indeed, the table shows that customer queries and product returns had a strong negative (or inverse) correlation of -0.8908. This means more queries resulted in less returns. Presumably, fewer returns lead to more sales. We found there to be an inverse correlation of -0.7159. A natural assumption would be that customer queries would therefore, by the transitive property, have a positive correlation with sales. Indeed, we found that the correlation was very high at 0.8211. Customer queries had a lower positive correlation to total competitor market shares and employee count at 0.6207 and 0.6347, respectively.

The chart breaking down the sales over years by total and individual product does not seem to show a trend on a per product basis.

A closer look at correlation calculations shows that Product A correlates most closely with sales and the correlation is high at 0.7277.

|  |  |  |  |
| --- | --- | --- | --- |
|  | *Sum of Product\_A\_Sales* | *Sum of Product\_B\_Sales* | *Sum of Product\_C\_Sales* |
| *Sum of Product\_A\_Sales* | 1 |  |  |
| *Sum of Product\_B\_Sales* | 0.146998636 | 1 |  |
| *Sum of Product\_C\_Sales* | 0.041424693 | -0.39402994 | 1 |
| *Sum of Total Sales* | 0.727652481 | 0.360290099 | 0.507900675 |

As mentioned in the previous section the decline in overall expenses appears to be due to the decline in operating expenses, specifically. Indeed, the correlation between the two is a strong 0.7897.

Another strong correlation is total sales and employee count at 0.7483. On topic, customer queries and employee count have a lower but still relevant positive correlation of 0.6347. Additionally and also relevantly, employee count and product returns have a very strong inverse correlation at -0.7935.

Other positive correlations that are noted here but will be left alone are sales and competitor market share (0.7694), expenses and industry growth (0.7433), and competitor market share and employee count (0.9830).

Other inverse correlations that are noted here but will be left alone are the year and GreenTech Innovations market shares (-0.7279), the year and industry growth (-0.9375), competitor market shares and returns (-0.7163), and queries and GreenTech Innovation market shares (-0.7274).

Revenue, sales, and expenses cannot be accurately predicted for 2024 because they do not have a strong linear trendline (R-Squared values 0.1288, 0.012, and 0.3514, respectively). The best trendline we have is industry growth rate at 0.8789 which is steadily declining from 2019 to 2023. Industry growth rate has a fairly strong positive correlation with expenses (0.7433) and a somewhat strong inverse correlation with queries (-0.4510). These are both positive indicators for GreenTech Innovations as if this trend continues, expenses might continue to fall while queries, related to returns and then revenue, will rise.

This should be met with caution, however. GreenTech Innovations’ total expenses are driven by operational expenses which are at a historic low for the five year period. It could be a bad idea to assume that expenses will continue to fall in the future regardless of whether or not industry growth continues its downward trend.

**Probability and Decision Frameworks**

Based on the analysis so far, we have found that queries/returns, employees, and Product A are closely correlated with sales. We feel that the overall economy is stable and do not see reason to anticipate massive growth or massive recession. Most of our trendlines related to sales are relatively flat so we do not recommend pursuing either the maximax or maximin decision-making framework.

While the maximax decision-making framework offers the highest potential payoff, it is best suited for optimistic or risk-seeking individuals. This approach is best suited for an optimistic market. It focuses on choosing the alternative that has the maximum possible gain, assuming the best possible outcome for each option. Essentially, it involves evaluating the potential maximum payoff for each decision and selecting the option with the highest of these maximum payoffs. As our trendlines over the last five years are mostly flat we caution against this strategy.

The maximin is the opposite of maximax. Instead of focusing on maximizing potential gains, maximin focuses on mitigating potential losses. It involves evaluating the worst possible outcome for each decision alternative and then choosing the option with the best worst-case scenario. Essentially, this approach prioritizes the option that has the least severe negative impact if things don't go as planned. It is often employed by decision-makers who are risk-averse and seek to safeguard against the most adverse outcomes, ensuring that the worst-case situation is as favorable as possible.

Instead, we recommend the realism decision-making framework. Also known as the Hurwicz criterion, a weighted average is calculated by multiplying the impact of an outcome by the associated probability of the outcome. This is a balance of the aggressive maximax approach and the defensive maximin approach.

The probability model given by GreenTech Innovation is as follows:

|  |  |
| --- | --- |
| **Market Scenario** | **Probability** |
| Economic Boom | 0.3 |
| Stable Economy | 0.5 |
| Recession | 0.2 |

To help illustrate a decision framework, we will use 2023 data for queries, employees, and Product A as the baseline assigned to a stable economy. For an economic boom, we will add 10%. For a recession scenario, we will subtract 10%. The expanded table with the three major variables contributing to sales is as follows:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Market Scenario** | **Probability** | **Queries** | **Employees** | **Product A** |
| Economic Boom | 0.3 | 410.3 | 114.4 | $67,412.40 |
| Stable Economy | 0.5 | 373 | 104 | $61,284.00 |
| Recession | 0.2 | 335.7 | 93.6 | $55,155.60 |

The Expected Monetary Value (EMV) of Product A is straightforward and follows the standard expected monetary value formula:

EMV = Impact of Occurrence \* Probability of Occurrence

In the case of Product A:

($67,412.40 \* 0.3) + ($61,284.00 \* 0.5) + ($55,155.60 \* 0.2) = $61,896.84

While the queries and employees are not monetary values, the same principle can be applied to find the expected count or expected count value in each of the ‘states of nature’ or uncontrollable, potential future conditions or outcomes.

For queries, we have:

(410.3 \* 0.3) \* (373 \* 0.5) + (335.7 \* 0.2) = 376.73

For employees, we have:

(114.4 \* 0.3) + (104 \* 0.5) + (93.6 \* 0.2) = 105.04

In summary:

|  |  |
| --- | --- |
| EMV Queries | 376.73 |
| EMV Employees | 105.04 |
| EMV Product A | $61,896.84 |

We see that each of these values is higher than the current 2023 values but only barely. This suggests that preparing for a best case scenario, the maximax approach, or preparing for the worst case scenario, the maximin approach, probably are not necessary. A balance between the two, criterion of realism, is likely a better choice.

**Stakeholder Feedback**

The customers clearly value product quality and usability. Product quality is crucial because it directly impacts customer satisfaction, brand reputation, and overall business success. High-quality products meet or exceed customer expectations, leading to increased trust, loyalty, and positive word-of-mouth, which can drive repeat business and attract new customers. Consistent quality helps differentiate a company from competitors, often allowing for premium pricing and market leadership. Moreover, it reduces the likelihood of defects and returns, which can save costs and enhance operational efficiency. Ultimately, maintaining high product quality is essential for long-term profitability and sustainable growth.

Product usability is vital because it directly affects how effectively and efficiently users can interact with a product. A product with high usability ensures that users can easily understand and operate it, leading to a better user experience, increased satisfaction, and reduced frustration. Good usability often translates into fewer errors, lower training and support costs, and higher productivity, as users can complete tasks more quickly and accurately. Additionally, products that are easy to use are more likely to achieve higher adoption rates and positive reviews, which can enhance a brand's reputation and competitiveness in the market.

Because Product A is the most closely correlated with sales it stands to reason that Product A has superb quality and usability. This criteria should be used as a primary metric by which to evaluate products at GreenTech Innovations. Products B and C should be reevaluated to see how they might be able to be made higher quality or more usable. Positive responses to customer service, innovation, and pricing are encouraging and these metrics should be given secondary priority.

Negative responses are highest in quality and innovation. Product innovation is crucial because it drives competitive advantage, meets evolving customer needs, and stimulates growth. By introducing new or improved products, companies can differentiate themselves from competitors, capture market share, and attract new customers. Innovation helps businesses adapt to changing market conditions, technological advancements, and consumer preferences, ensuring they remain relevant and responsive. Additionally, innovative products can open up new revenue streams and enhance brand reputation, leading to increased customer loyalty and long-term success. Overall, product innovation is essential for maintaining a dynamic, forward-thinking organization in a rapidly evolving market. For green companies, product innovation is especially important because it aligns with their commitment to sustainability and environmental responsibility. Innovative products can help reduce environmental impact by incorporating eco-friendly materials, energy-efficient technologies, and waste-minimizing processes. This not only supports the company's mission to promote environmental stewardship but also differentiates it in a market increasingly driven by sustainability concerns. Additionally, innovation in green products can meet growing consumer demand for environmentally conscious options, enhance brand reputation, and comply with evolving regulations and standards. Overall, product innovation is crucial for green companies to achieve their sustainability goals while driving growth and maintaining a competitive edge.

**Conclusion**

After reviewing GreenTech Innovation’s data, we have several recommendations. To increase profitability, GreenTech Innovation should first focus on those variables that are most closely related to sales. Customer queries correlates to fewer returns which correlates to increased sales. Customers contacting employees directly or via customer service leads to fewer returns and ultimately increased sales. We note that customer queries have a positive correlation with employee count. Additionally, we found that there is a direct correlation between employees and sales. Our first recommendation, therefore, is to retain and hire more employees and have more employees on hand during working hours.

Our second finding is that Product A is most associated with high sales. The survey feedback helped us understand exactly which aspects of Product A make it desirable and a high seller. Namely, quality and usability. Our second recommendation, therefore, is to try to improve Product B and C’s quality and usability and also to focus on these attributes when developing future products.

Our recommendation for a decision-making framework is the balanced criterion of realism. We do not see a strong reason to pursue an overly aggressive or overly defensive decision-making approach in the current competitive environment given the data we have seen.

**References**

Albright, S. C., & Winston, W. L. (2019). Business Analytics (7th ed.).

Ehling, P., & Heyerdahl-Larsen, C. (2017). Correlations. *Management Science*, *63*(6), 1919–1937. https://doi.org/10.1287/mnsc.2015.2413

Zhang, X., & Hedeker, D. (2022). Defining R‐squared measures for mixed‐effects location scale models. *Statistics in Medicine*, *41*(22), 4467–4483. <https://doi.org/10.1002/sim.9521>

Mongin, P., & Pivato, M. (2021). Rawls’s difference principle and maximin rule of allocation: a new analysis. *Economic Theory*, *71*(4), 1499–1525. https://doi.org/10.1007/s00199-021-01344-x

Bleichrodt, H., Grant, S., & Yang, J. (2023). Testing Hurwicz Expected Utility. *Econometrica*, *91*(4), 1393–1416. https://doi.org/10.3982/ECTA19221