Predicting Catalog Demand

REVIEW | HISTORY

Meets Specifications

Hello Udacian!
Congratulations!!! You made it. 🎉 😄
You should be proud of your work, I know I am proud of it.
This is your first submission and it is really outstanding. :+1:
I am sure you have enjoyed working on this project, I can see the hard work in your project.
I really hope the lectures and this project in particular have effectively taught you the fundamentals of this course.
Well Done 👏
It was a pleasure reviewing your project.
I wish you all the best!
Happy Learning 😇

Business and Data Understanding

✓ The section is written clearly and is concise. The section is written in less than 500 words.

Good job!
Your answers are clear and easy to understand, it is not more than 500 words. 👏
Good job writing a brief and to the point answers.

✓ All following questions have been accurately answered:

- What decisions need to be made?
- What data is needed to inform those decisions?

Good job! You have answered all questions correctly.

> **What decisions need to be made?**
> Whether to send the catalog to those 250 customers ? i.e. catalogs will only be sent if the expected profits exceed $10,000.
> **What data is needed to inform those decisions?**
> Basically we need to predict sales and profit, so the data to predict sales and expected profit are Cost of Catalog, Customer Segment, , _ScoreYes, Average Number of Product Purchased and Margin.

Analysis, Modeling, and Validation

✓ The section is written clearly and is concise. The section is written in less than 500 words.

Good job writing a brief and to the point answers. 👏

✓ Each predictor variable is shown to have a linear relationship between the target variable whenever appropriate.

Each predictor variable should be significant (p-value <= 0.05)

Good job!
Your linear model shows the correct values for R-squared and P-values (p-value <= 0.05) and you have shortlisted the correct predictor variables for the linear model. Well done!
You have created the scatter plot as well to show the existence of a linear relationship between continuous predictor variables and the target variable.
`Suggestions:`

> **Correlation coefficient and p-values:**
>
> - The correlation coefficient, r, tells us about the strength and direction of the linear relationship between x and y. The correlation coefficient is a number between –1 and 1 that determines whether two paired sets of data related. The closer to 1 the more 'confident' we are of a positive linear correlation and the closer to –1 the more confident we are of a negative linear correlation.
> - The P-value is the probability that you would have found the current result if the correlation coefficient were in fact zero (null hypothesis). If this probability is lower than the conventional 5% (P<0.05) the correlation coefficient is called statistically significant.

You can also checkout this part of the course for the topic "Multiple Linear Regression with Excel":
https://classroom.udacity.com/courses/ud976/lessons/4e33b70a-72a4-47cb-959a-28632ae6aaff/concepts/631d190c-8626-4dd7-92df-f5bd96913c48

✓ p-values and R-squared values are used to justify how well the linear model works.

To judge the goodness of the linear model you should be looking at the report generated from the R-squared value of the Linear regression tool.
You have used the value of p-values and R-squared values that you got from the report for Linear model to Explain why you believe your linear model is a good model. Well Done. 👏🏻

> - The higher the `adjusted r squared value`, the higher the explanatory power of the model. This value represents the amount of variation in the target variable explained by the variation in the predictor variables. Any model with an adjusted r-square value above 0.70 is considered to be a strong model. Our present linear model has a value of `0.84`, hence it is a good model.
> - In the previous section , I have explained the concept of p-value in detail i.e. when a predictor variable has a `p-value` below `0.05`, we can conclude that the relationship between it and the target variable is considered to be statistically significant.
> In this case it is much lesser than 0.05 threshold. Therefore each predictor variable is significant to the linear model.

✓ The regression equation given is correct. Each coefficient should have up to 2 digits after the decimal figures (ex: 1.28).

Well done!
The regression equation given is Perfect. 👏🏻

Presentation/Visualization

✓ The section is written clearly and is concise. The section is written in less than 500 words.

Good job!
Your answers are clear and easy to understand. 👏🏻

✓ All questions have been accurately answered and the recommendations are well justified.

- What is your recommendation?
- How did you come up with your recommendation?
- What is the expected profit from the new catalog (assuming the catalog is sent to these 250 customers)?

Thank you for a clear description of the entire project and steps involved in drawing the conclusion. Well done!
Your answers are correct for :

> - **What is your recommendation?**
> The expected profit is $21,987, which is higher than $10,000, so the company should send the catalog to these 250 new customers.
> - **How did you come up with your recommendation?**
> You have used a data-driven approach to justify your answers. 👏🏻
> - **What is the expected profit from the new catalog (assuming the catalog is sent to these 250 customers)?**
> The expected profit is $21,987.

✓ The profit calculation is correct.

Well done!
The expected profit from the new catalog is `$21,987.44`. :+1:

DOWNLOAD PROJECT

RETURN TO PATH