1. Task Description

Team,

Mike Dobson, our CEO, heard about the results of our data mining efforts and would like to understand them better. He is considering funding an expansion of our data mining efforts.

To that end, I would like you to prepare and deliver a presentation that explains the questions you have helped the sales team answer, the methods you used to answer those questions, and the results you delivered. In addition, I would like to get Mike thinking about how we can leverage data mining in the future to make data-driven decisions about all aspects of our business. Would you please conduct some research and create a slide that lists the types of questions data mining can help to answer; one to two slides should be sufficient.

Remember, Mike doesn't have much data mining knowledge beyond a few articles he has read. Please make sure that the presentation is appropriate.

In addition to the presentation for the CEO, I'd also like you to create a two- or three-page report detailing key take-aways and lessons you have learned about data analytics, making specific recommendations for how Blackwell should approach such projects in the future. This report will just be seen internally by me so you can make it as technical as you feel it needs to be.

Thanks,

Danielle

Danielle Sherman
Chief Technology Officer
Blackwell Electronics
www.blackwellelectronics.com

1. Task Solution

The solution for this task is divided in two documents, both documents will be loaded in the classroom.

- Task 5A – Data Mining Slides Deck.pdf

- Task 5B - Lessons Learned Report.docx

The lessons learned report will address problems, impact and recommendations for the following categories:

- Category 1 - Get started with R

- Category 2 – Classification: Predict which Brand of Products Customers Prefer

- Category 3 - Multiple Regressions

- Category 4 - Discover associations between products

- Category 5 – General

Table 1 - Lessons Learned

Category	Topic	Problem / Opportunity	Impact	Recommendation
1, 2, 3 & 4	Resource Section	Outdated resources / Value added resources	High – Outdated resources require additional time to search for the right resources.	Provide operational links and value-added resources
	Plan of Attack	Use of plan of attack	High – The plan of attack was not helpful in this module, since outdated resources and low value-added link were provided (resource section).	Include a Datacamp subscription in the program since other programs includes this in their plan to leverage coding knowledge between the students, this works for R and Python as well

**Task 5: Present Data Mining Activities to CEO
and Lessons Learned Report**

Student: Steven Melendez Lara – Group: 6-C

Category	Topic	Problem / Opportunity	Impact	Recommendation
	R	Developed expertise in a very common programing language	High – The R programing language is widely used, the knowledge of this tool will distinct us between other candidates in job field	Buy a Datacamp subscription since the resources of this course are very limited to understand R coding and capabilities
	Results interpretation	Results discussion section	High – Results discussion and recommendations are key and important to provide technical and business criteria to the problems	Include results discussion in each task to easily summarize results. These summaries help a lot during the slides deck creation.
2 & 3	Machine Learning	When to use each machine learning algorithm	Medium – Difficulty to define the use of each algorithm to complete the tasks	Provide examples of thoughts processes and rationales to use the different algorithm for each situation
All	Collaboration	Peer Review	High – The collaboration with other team members helps to validate your work and also to detect potential errors	Always conduct a peer review with at least with one classmate
4	Market Basket Analysis	Resource Section /	Medium – MBA was not easy to understand and implement, the resource section lacks of a detailed example to pick the idea / concept of the MBA	Update resource section with other links with examples such as: https://www.datacamp.com/community/tutorials/market-basket-analysis-r
General	Mentorship	Time Management	High – The use of mentorship time was improved and now the discussions are more productive	None, Keep doing this way