PROJECT – DAL371

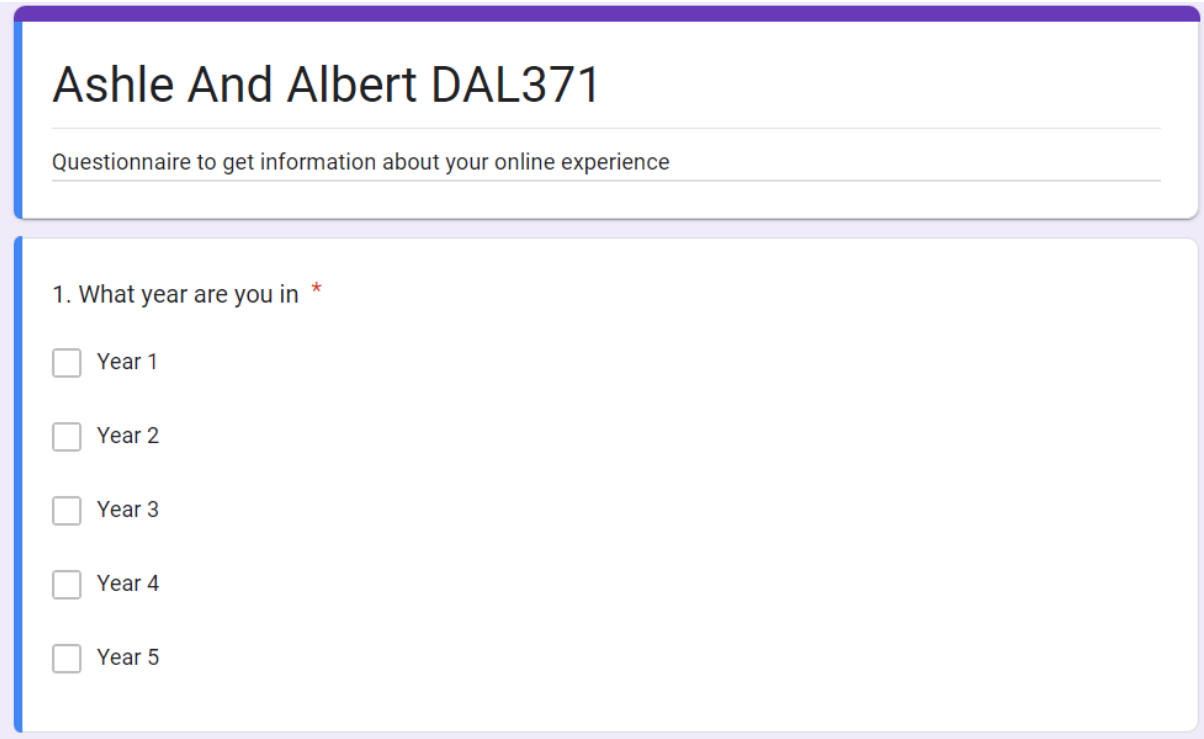
CREATORS
ASHLE OLIVIER 578325
ALBERT WOLFAARDT 575709

Table of Contents

Data Sources

1. When incorporating datasets such as Harvard Dataverse (Pulickamadhom Sreedhar, 2021), UC Irvine Machine Learning Repository (Pulickamadhom Sreedhar, 2021), and Kaggle (SUZAN, 2022), we utilize custom off-the-shelf data and external data sources. This category of data refers to freely accessible, pre-structured data which we will modify to align with the requirements of our project.
2. Content: To gather insights for enhancing our existing online system or potentially adopting a new one, we designed a survey and gathered feedback from numerous third-year students regarding their experiences with Hybeflex and other online classroom platforms. This valuable information allows us to identify areas for improvement and align our system with the preferences and requirements expressed by the students.

The following are the questions included in the questionnaire. It's important to note that while not all questions will be utilized for this specific project, we aimed to gather a comprehensive understanding of the students' responses, which can be beneficial for future projects as well.



Ashle And Albert DAL371

Questionnaire to get information about your online experience

1. What year are you in *

- [ ] Year 1
- [ ] Year 2
- [ ] Year 3
- [ ] Year 4
- [ ] Year 5

2. What course are you studying *

- [ ] Bachelor of computing
- [ ] Bachelor of Information Technology
- [ ] Diploma in Information Technology
- [ ] Diploma for Deaf Students
- [ ] Certificate: IT
- [ ] National Certificate: IT

3. Do you prefer studying using online or in person classes *

- [ ] Online
- [ ] In person

4. Has your experience with the current online platform been enjoyable *

- [ ] True
- [ ] False

5. What platform would you choose if the school decided not to use Hybeflex *

- [ ] Teams
- [ ] Microsoft Classrooms
- [ ] Discord
- [ ] Zoom
- [ ] Other...

5.1. What makes you enjoy this specific platform, chosen in Question 5, compared to Hybeflex *

Long answer text

6. Do you feel that lecturers interact enough with the online students *

- [ ] True
- [ ] False

7. What do you like about the current existing system

Long answer text

8. What do you dislike about the current existing system

Long answer text

9. Do you have any suggestions for the existing system

Long answer text

Transformations that will be carried out.

To facilitate the viewing of the transformations, we have included both the original files and the Power BI document where most of the transformations were performed.

1. We will be changing the Excel data files into .CSV format. This is because .CSV files are more suitable to the Power BI platform as they are already able to be made into tables instead of having to find the data in Excel and convert it to a table. The run time for Excel files is also very slow compared to .CSV files (Webb, 2018). Furthermore, if we wish to send this file with the report it is easiest to send it in .CSV format as this format can be opened in many different applications (Lahar, 2020). Therefore, if the person who is sent this data doesn't have access to Excel, they can open it on any text editor (Lahar, 2020).
2. Transformations carried out with Ashle And Albert DAL371 (Responses) file:
   2.1. The CSV file provided comprises all the data collected from the distributed questionnaire.
   2.2. After careful consideration, we have opted to eliminate the first two columns from the dataset. The first column, which contained the timestamp, is unnecessary since all the information was collected within a two-week timeframe. Likewise, the second column indicating the current year of study is redundant as the questionnaire was exclusively sent to third-year students.
   2.3. To provide clearer comprehension of the column contents, we renamed them to better reflect the full questions they represented.
   2.4. Furthermore, we made the decision to retain only the columns pertaining to questions 2, 3, 4, 5, and 6, as these columns are most suitable for visualization purposes. The remaining questions can be utilized in future projects as needed.
   2.5. Due to the inclusion of both "true" and "false" responses for question 4 and 6 by certain participants, which is not in line with our requirements, we have removed the rows associated with these respondents from the dataset.
   2.6. Our focus for question 5 was on the initial options selected by the students, as these choices likely represent their primary preferences if given the opportunity to switch to a different platform.
   2.7. Given that columns 3 and 5 exclusively consist of true or false values, we have converted their data type to Boolean (True/False) for consistency and ease of interpretation.
   2.8. To be able to use the data in our visualizations, we will create an index for each row.
3. Transformations carried out with C_Irvine_Machine_learning_repository_data file:
   3.1. The web page associated with this document (https://archive-beta.ics.uci.edu/dataset/697/predict+students+dropout+and+academic+success) provides explanations for the meaning and representation of each number mentioned in the dataset.
   3.2. This dataset was created in 2021.
   3.3. We have narrowed down our focus to specific columns of interest, which include:
      3.3.1. Course: In this analysis, we will exclusively consider students who are enrolled in Informatics Engineering (IT), which is the specific program offered by our university for students to pursue their studies. (9119 according to the website)
      3.3.2. Curricular units 1st sem (grade): We will examine the grades of students in their curricular units during the first semester to assess their academic performance.
      3.3.3. Curricular units 2nd sem (grade): This column provides information about the grades of students in their curricular units during the second semester, further aiding our assessment of their academic performance.

3.4. We proceeded to replace the full stops in column 2 and column 3 with commas, enabling us to modify the data type to decimal numbers. This adjustment facilitates better compatibility and interpretation of the data.

3.5. We have incorporated a custom column using DAX (Data Analysis Expressions) to calculate the average performance for each semester. This allows us to assess the average academic performance of different students throughout the year.

4. Transformations carried out with Student and Educator Perceptions of Online Classes - Base data file:

4.1. We have carefully selected the following columns for specific reasons, while removing other columns from consideration:

4.1.1. Are you a student or an educator?: This column provides insights into educators' perspectives on online platforms and their observations regarding how other educators have managed aspects such as communication. Additionally, it allows us to explore students' sentiments and opinions on the topics represented by the following columns.

4.1.2. How do you rate yourself on your ability to use virtual meeting platforms, on a scale of 1-10?: This column helps gauge students' self-perceived proficiency in using virtual meeting platforms, providing insights into their readiness for online learning.

4.1.3. How would you rate your internet connectivity on a scale of 1-10?: This column assists in assessing students' internet connectivity quality, ensuring that those who envision using an online learning platform have the necessary resources.

4.1.4. How effective do you think online classes are in communicating content and intent in a classroom, on a scale of 1-10?: This column provides insights into students' perception of the effectiveness of online classes in terms of content and intent communication, helping identify potential areas of improvement.

4.1.5. Students, how much effort do you think is put by educators to facilitate online learning as compared to regular sessions at school? Educators, how much effort do you think students put into education as opposed to regular sessions at school? On a scale of 1-10?: This column allows us to assess students' perceptions of educators' effort in facilitating online learning and educators' perceptions of students' engagement, assisting in making informed decisions regarding teaching methodologies and support.

4.1.6. How would you rate your ability to concentrate during online classes, on a scale of 1-10?: This column provides insights into students' ability to concentrate during online classes, helping establish strategies for maintaining engagement and scheduling appropriate break times.

4.1.7. Do long hours in a virtual classroom cause discomfort such as headaches, sleep disturbances, back pain?: This column enables us to understand whether prolonged virtual classroom sessions lead to discomfort and physical issues, aiding in determining optimal class durations for online learning.

4.2. We will then rename the columns to make them more suitable to use in visualizations.

4.3. We have included an index column in the dataset, which allows us to view each entry separately and uniquely identify them. This index column serves as a sequential numbering system that facilitates the organization and referencing of individual records within the dataset. By having an index column, we can easily track and analyse specific data points or perform targeted operations on specific entries as needed.

5. Transformations carried out with students_adaptability_level_online_education file:

5.1. Using first column as headers.

5.2. Based on our preferences, we have identified the following columns to be used in the analysis and have removed the rest:

5.2.1. Education level: We will focus on university students as Belgium Campus is an institute that specifically caters to this level of education.

5.2.2. Institution type: We will only consider students from non-government institutions, as Belgium Campus is also a non-government university.

5.2.3. IT student: We will specifically examine IT students as that is the field of study available at Belgium Campus.

5.2.4. Loadshedding: Considering the power outages in South Africa, we want to understand how loadshedding impacts students' adaptability.

5.2.5. Internet Type and Network Type: We aim to assess if certain internet and network types are necessary for students to have to achieve a high pass rate when using the platform.

5.2.6. Class duration: This column will enable us to examine if the duration of classes influences students' adaptability.

5.2.7. Adaptivity level: This column will serve as the reference point against which the other columns will be compared.

Charts that will be used and why.

We made the decision to generate two reports: one utilizing the fundamental Power BI desktop, and the other employing an R script integrated with Power BI. The initial report presents visualizations that assess the feasibility of either updating our existing system or adopting a new online learning platform for Belgium Campus. Meanwhile, the second report highlights the potential enhancements that Belgium Campus can incorporate into the system to maximize its effectiveness. You can refer to the enclosed Power BI file to view all the charts and corresponding code.

First Report

1. We have chosen to utilize a Line Chart to compare the adaptability reported by students in the Kaggle Dataset with the corresponding level of load shedding. Through this analysis, we have observed a relationship where, as the load shedding status increases to a higher level, fewer students are able to adapt to using an online learning platform effectively.
   This finding is significant and should be carefully considered by the executives at Belgium Campus. It highlights the potential impact of load shedding on students' ability to actively participate in online learning, whether through improving the existing platform or introducing a new one. The executives may need to revaluate their decision regarding the implementation of the project, considering the potential challenges posed by load shedding and its influence on students' ability to engage with the platform effectively.
2. We have utilized a Gauge Chart to analyse the average scores within the range of 0 to 20, derived from the total average calculated by the DAX column in the C_Irvine_Machine_learning_repository dataset. This analysis provides insights into the performance of students in the year 2021, specifically in relation to their overall average scores. The gauge chart indicates that the average score for students in IT was 8.68, which suggests that most students did not even achieve a 50% pass mark for their academic performance throughout the year. This finding highlights the challenges faced by students in studying IT and raises important considerations for the executives at Belgium Campus.
   The executives need to carefully evaluate whether the implementation of an online learning platform would positively impact these average scores or potentially have a detrimental effect. This data serves as an indicator for them to reflect on the potential benefits and drawbacks of an online platform in improving the academic performance of students studying IT.
3. To gauge student preferences between online and in-person classes, we have created a column chart. This chart is significant for Belgium Campus as it helps assess the level of interest in utilizing an online platform. It allows the institution to determine whether investing time and resources into developing a new system or modifying the current one is justified, based on the level of student interest.
   The column chart was constructed by combining the courses indicated by students as the y-axis, while the x-axis represents the choice between online and in-person classes. This visual representation provides a clear overview of students' preferences and aids in making informed decisions regarding the future of the learning platform at Belgium Campus.
4. To assess the feasibility of virtual classes for both students and educators, we have created a scatter plot using the Student and Educator Perceptions of Online Classes - Base dataset. The scatter plot incorporates the index of the dataset as well as the ratings provided by students regarding their ability to use virtual meeting platforms.
   This analysis is crucial in determining whether the implementation of an online learning platform is viable. By examining the scatter plot, we can observe the distribution of ratings given by students. Notably, most students have rated their ability to use virtual meeting platforms at 6 and above on a scale of 1 to 10.
   This finding indicates that a significant number of students possess the necessary skills and proficiency to actively participate in online classes. Consequently, the creation or improvement

of an online learning platform holds potential feasibility, as a substantial portion of the student population demonstrates a positive outlook towards online classes. Upon examining the perspective of teachers, we observe that most of them have provided ratings of 8 and above in terms of their ability to participate in online classes. This finding further supports the viability of creating or upgrading the online learning platform.

The high ratings from teachers indicate their confidence and proficiency in utilizing virtual meeting platforms for conducting online classes.

Considering both the positive ratings from students and teachers, it becomes evident that the implementation or enhancement of an online learning platform is highly feasible. The collective readiness and ability of both students and educators contribute to the overall viability of the project, providing a strong foundation for successful adoption and utilization of the online platform.

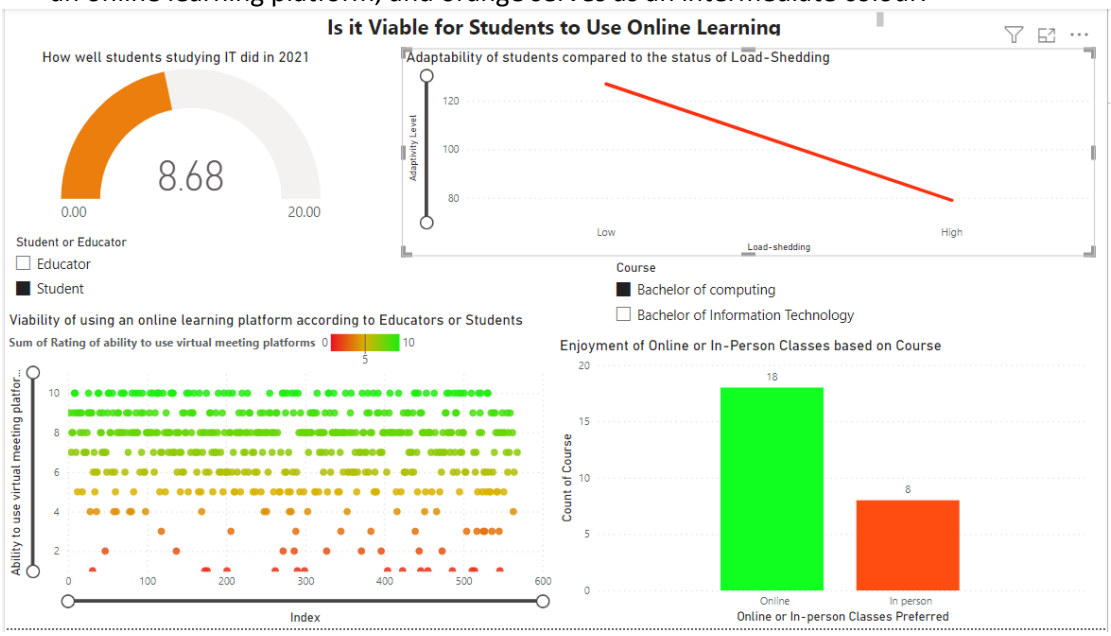
5. To enhance the scatter plot visualization, we included a slicer that allows executives to toggle between the responses of educators and students. This provides a convenient way to compare and analyse the perspectives of both groups.
6. We have incorporated a slicer into the bar chart visualization to further analyse the perceptions of different degree programs regarding online and in-person classes. This slicer allows users to selectively view the perspectives of each degree type, enabling a comprehensive comparison of their preferences.

   By leveraging this slicer, stakeholders can easily identify the varying sentiments of different degree programs towards online and in-person classes. The visualization indicates that across all degree types, students generally express a greater preference for online classes as compared to in-person classes.

   This consensus among students, regardless of their degree program, suggests a positive reception towards online learning at Belgium Campus. It indicates that implementing an online platform aligns with the preferences and satisfaction of the student body, affirming the feasibility and value of adopting such a learning environment.

   The presence of the slicer enhances the flexibility and interactivity of the bar chart, facilitating detailed exploration of the data based on specific degree programs and their corresponding opinions on the mode of instruction.

   We have chosen a colour scheme consisting of red, orange, and green for visual indicators. In this scheme, red represents non-viable levels or values, green represents high viability for using an online learning platform, and orange serves as an intermediate colour.



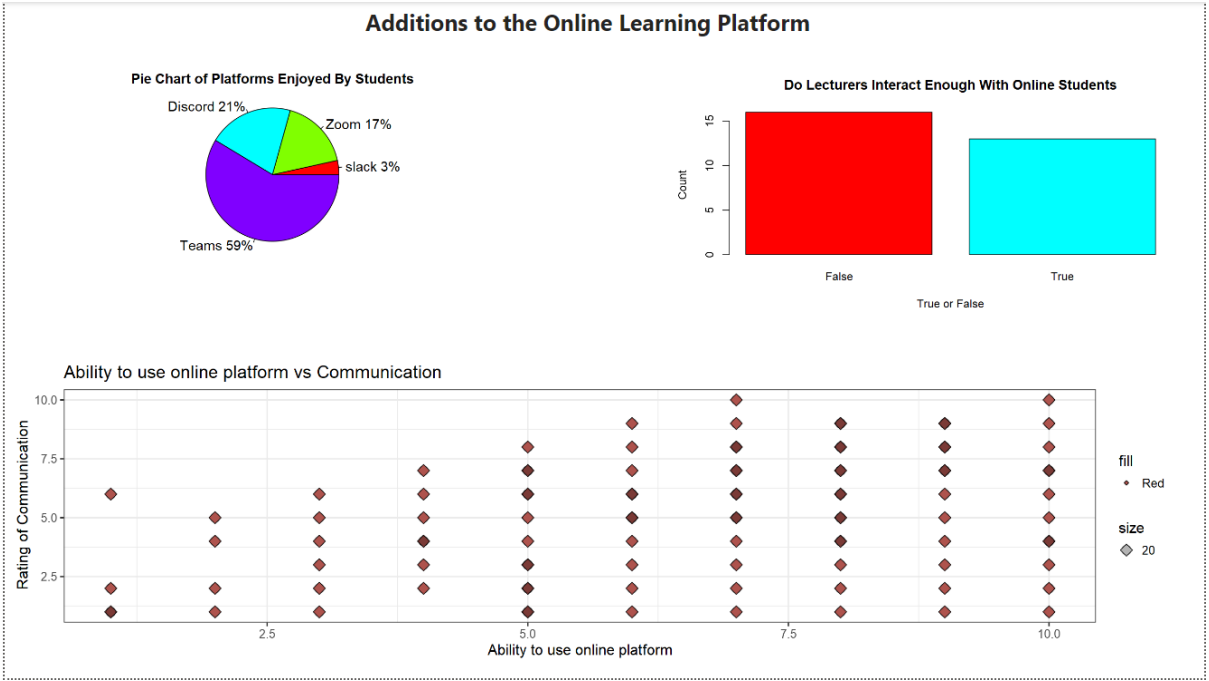
Second Report

1. We made the decision to employ a pie chart to visually convey the preferences of students regarding existing platforms. This information will assist the platform designers in conducting research on the most popular platforms and incorporating their desirable features into their own creation. Based on the pie chart data, it is evident that the most favoured platform is Teams. Hence, the programmers should concentrate their research efforts on identifying and incorporating the features present in the Teams platform.
2. For the second visualization, we aimed to illustrate the correlation between the proficiency of using an online platform and the effectiveness of communication by educators in a classroom setting. To achieve this, we utilized a scatter plot, which indicated a positive relationship between improved communication and enhanced ability to utilize the platform. Consequently, based on these findings, we strongly recommend that Belgium Campus consider sending their lecturers on a course specifically designed to enhance their communication skills when utilizing the online learning platform.

   To focus solely on student information, we eliminated rows from the dataset that contained the term "educator" in the Student or Educator column.
3. After assessing the level of interaction between lecturers and online students at Belgium Campus, we proceeded to create a bar plot. This visualization aimed to gauge the students' perception of whether the lecturers adequately engage with online students. This includes aspects such as regularly checking messages, ensuring online students are following the class, and sharing relevant links through the online platform's chat feature.

   Based on the findings from the bar plot, it became evident that lecturers are not sufficiently engaged with online students. Despite the existence of lecturer evaluations at Belgium Campus, these evaluations have not resulted in noticeable improvements in lecturer interaction with online students.

   As a recommendation, we suggest that the executives consider providing lecturers with two monitors. One monitor can be dedicated to screen sharing, while the other can be used specifically to monitor and respond to online chats. This setup would facilitate better multitasking for lecturers and enable them to effectively engage with online students during live classes.



References

Lahar, S., 2020. *What Is a CSV File? Guide to Uses and Benefits.* [Online]
Available at: https://flatfile.com/blog/what-is-a-csv-file-guide-to-uses-and-benefits/
[Accessed 10 May 2023].

Pulickamadhom Sreedhar, S., 2021. *Harvard Dataverse, Student and Educator Perceptions of Online Classes, V1.* [Online]
Available at: https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/2AAVSX
[Accessed 05 2023].

SUZAN, M. M. H., 2022. *Students Adaptability Level in Online Education.* [Online]
Available at: https://www.kaggle.com/datasets?search=online+learning
[Accessed 8 May 2023].

The World Bank Project, 2020. *The World Bank.* [Online]
Available at: https://data.worldbank.org/topic/education?locations=ZA
[Accessed 10 05 2023].

Webb, C., 2018. *Comparing The Performance Of CSV And Excel Data Sources In Power Query.* [Online]
Available at: https://blog.crossjoin.co.uk/2018/08/02/comparing-the-performance-of-csv-and-excel-data-sources-in-power-query/
[Accessed 10 May 2023].