Criteria (learning outcome)	Domain	0 - Incomplete	1- Fair	2- Good	3- Excellent	Assessment event in which applies
Identify business problems and opportunities for improvement through data analysis, formulate a well-defined hypothesis, research question or problem statement, and apply appropriate data analysis techniques to answer the question.	**Business Problem Solving**	**All of these apply:** • There is no clear hypothesis or research question defined • The data analysis techniques applied are inadequate for data and business problem • Doesn't provide clear or relevant insights from analysis	**At least two of the following apply:** • There is **one well** defined hypothesis or research question and this is not too broad or narrow, analytical and neither too complex or simple • The data analysis techniques applied may be **somewhat effective**, but there is room for improvement. • Provides **clear** insights from analysis to aid decision-making	**At least two of the following apply:** • Includes **at least two** clear and **focused** hypothesis or research questions • Demonstrates a **solid** understanding of data analysis principles by utilizing **appropriate** techniques that match the data and business problem, with clear rationale for their selection • Provides **clear** and **relevant** insights from analysis to aid decision-making	**At least three of the following apply:** • Identifies **multiple relevant** business problems and opportunities for improvement through data analysis, with **clearly** defined hypothesis or research questions • Demonstrates an **exceptional** understanding of data analysis principles by utilizing relevant techniques that match the data and business problem, with **outstanding** rationale for their selection • Delivers **clear**, **impactful** insights that facilitate decision-making by translating data analysis results into actionable recommendations with a **substantial** impact on the business problem or need • Demonstrates an **exceptional** commitment to driving continuous improvement and innovation through data analysis	Project 1 (DW) Project 2 (EDA) Project 3 (Final Project) Mini Project Web Scraping and APIs Mini Project SQL Kaggle Competition
Identify data quality issues in a dataset and apply appropriate data cleaning, wrangling, and manipulation techniques to address them.	**Data Preparation**	**All of these apply:** • No relevant data quality issues found • No relevant implementation of data cleaning or wrangling techniques	**All of these apply:** • **Some** data quality issues identified, but they are **irrelevant**, **unclear** or **incomplete** • Applied data cleaning/wrangling techniques, but these are **inadequate** for identified data quality issues	**At least three of the following apply:** • Identified **some relevant** data quality issues with **clear** explanations of their impact on analysis • Implemented **appropriate** data cleaning/wrangling techniques, **consistently** applied to most identified data quality issues • Used **some** data cleaning techniques but omitted handling null values and duplicates, dropping unnecessary columns, manipulating strings, or formatting data. • Addressed missing data, but **not fully** justified the imputation strategy	**All of these apply:** • Identified **most** relevant data quality issues with **clear** explanations of their impact on analysis • Implemented **relevant** data cleaning/wrangling techniques, **consistently** applied to all identified data quality issues • Used **many** data cleaning techniques including handling null values, duplicates, dropping unnecessary columns, manipulating strings, and formatting data • Addressed missing data **properly**, **fully justified** the imputation strategy	Project 1 (DW) Project 2 (EDA) Project 3 (Final Project) Mini Project Web Scraping and APIs Mini Project SQL Kaggle Competition
Apply Exploratory Data Analysis techniques to analyze data, validate hypotheses, draw conclusions and insights	**Data Analysis**	**All of these apply**: • No basic EDA techniques to analyze data, validate hypothese, draw conclusions and communicate findings • Improper use of EDA according to the data types	**At least two of the following apply:** • Basic EDA techniques but **incomplete** or **inaccurate** to analyze data, validate of hypotheses, draw conclusions and communicate findings • **Improper** or **incomplete** use of EDA techniques according to the data types • Use of charts and graphs using Python visualization libraries to present the data, but they may not be visually appealing or easy to understand • No use of inferential statistics, such as hypothesis testing with p-values, during correlation analysis to check for significant correlations or to check for normality (*)	**At least two of the following apply:** • **Competent** use of EDA techniques to analyze data, validate hypotheses, draw conclusions and communicate findings • **Competent** EDA techniques according to data types but could improve selection or their understanding of how results impact data analysis • Creates visually appealing charts and graphs using Python libraries, with room for improvement • No use of inferential statistics, such as hypothesis testing with p-values, during correlation analysis to check for significant correlations or to check for normality (*)	**All of these apply:** • Employs **sophisticated** EDA techniques to analyze data, validate hypotheses, draw data-driven conclusions and provide unique insights. • **Comprehensive** understanding of the data's characteristics, patterns, and relationships. Their proficiency encompasses thorough univariate, bivariate, and multivariate analysis, utilizing a range of numerical measures and graphical methods according to data type. **Outstanding** interpretation on how EDA results impact data analysis process (*) • Creates visually appealing and **highly informative** charts and graphs using Python libraries, showcasing a **remarkable** ability to communicate complex data in a clear and concise manner • Utilizes inferential statistics such as hypothesis testing with p-values to check for significant correlations during correlation analysis and to check for normality. Applies appropiate data transformations to ensure normality when necessary. (*)	Project 1 (DW) Project 2 (EDA) Project 3 (Final Project) Mini Project SQL Kaggle Competition * Inferential statistics and advanced EDA only applies for Project 2 and Project 3
Select and use appropriate data visualization techniques that effectively communicate insights, and create interactive and informative visualizations using Python libraries or visualization tools such as Tableau or Power BI.	**Data visualization and communication**	**At least two of these apply:** • No use of appropriate data visualization techniques that communicate insights • Dashboard lacks clarity, metrics and KPIs are not properly identified, well-defined, measurable or plotted (*) • Visual design is not informative	**At least two of these apply:** • **Fair** use of data visualization techniques to communicate insights, but needs improvement in some areas, like chart type selection and design refinement • Dashboard has **some clarity** but needs better organization (*) • KPIs and metrics need to be more clearly defined, measurable and plotted • Visual design needs improvement, some plots lack detail or have formatting issues	**At least three of these apply:** • **Solid** use of appropriate data visualization techniques that communicate insights effectively • Dashboard is **clear**, well-organized, visually appealing, and allows for decision-making (*) • KPIs and metrics are **clearly** define, measurable and plotted • Visualizations are **well-designed** with appropriate chart types and clear labeling of metrics	**All of these apply:** • **Exceptional** use of appropriate data visualization techniques that communicate insights effectively • Dashboard is **highly effective** with a clear and intuitive layout that conveys key insights at a glance (*) • KPIs and metrics are **expertly** defined**,** measured and plotted with **great attention** to detail • Visualizations are **expertly designed** with sophisticated use of chart types and creative visual elements	Project 1 (DW) Project 2 (EDA) Project 3 (Final Project) Mini Project SQL Kaggle Competition * Dashboard only mandatory for Project 2 + Final Project Case Study 2

Criteria (learning outcome)	Domain	0 - Incomplete	1- Fair	2- Good	3- Excellent	Assessment event in which applies
Develop proficiency in data analysis by applying rigorous methods for either predictive analysis using machine learning or descriptive analysis using statistics and SQL	**Data Analysis, Coding**	**Case Study 1 - all of these apply**: • No data preprocessing or inadequate to data and model selection • No application or inadequate selection of machine learning models • No or **inadequate** evaluation of model performance **Case Study 2** - **all of these apply**: • **Inadequate** basic SQL queries for data cleaning, manipulation, and analysis • No development of an appropiate database schema with proper organization and structure, which could potentially result in data inconsistencies • Doesn't combine SQL with Python for data analysis	**Case Study 1 - At least two of these apply:** • Attempts data preprocessing, but misses key issues that could impact model performance. • **Only one** machine learning model is applied and hyperparameter tuning is **not effectively** used • Appropriate metrics are not chosen to evaluate the model's performance • Satisfactory results are not achieved due to lack of important analysis steps **Case Study 2 - At least two of these apply:** • Use of basic SQL queries for data manipulation, but may miss some key concepts and more advanced queries are not developed • **Fair** development of an appropiate database schema. The data is **not properly** loaded into the database • Use SQL and Python together for data analysis, but may not produce highly meaningful insights	**Case Study 1 - At least three of these apply:** • Performs data preprocessing, addressing **most** issues. • **At least two** appropiate machine learning models are applied and hyperparameter tuning is used to optimize performance • Appropriate metrics are chosen to evaluate the model but **no appropriate** adjustments are made to improve model's performance • **Acceptable** results are achieved since most of important analysis steps were done **Case Study 2 - All of these apply:** • Use of basic SQL queries for data manipulation, but the student could improve more advanced queries like subqueries and window functions • Development of a **decent** database schema and **proper** loading of the data into the database • Use of SQL and Python together for data analysis with **meaningful** insights	**Case Study 1 - All of these apply:** • Performs and **justifies** data preprocessing, addressin **all** issues (why it's needed and the impact of doing it). If needed, inferential statistics and probability distributions are utilized to handle preprocessing steps (example: checking for normality). • **At least three** appropiate machine learning models are applied, ensuring the model chosen is suitable for the data (assumptions for the model are checked) and aligned with the business case. Hyperparameter tuning is used **proficiently** to optimize performance • Appropriate metrics are chosen to evaluate the model and appropriate adjustments are made to improve model's performance • **Outstanding** results are achieved since all of important analysis steps were done **Case Study 2 - All of these apply:** • **Proficient** use of basic and advanced SQL queries for data manipulation, such as subqueries and window functions. • Development of a **excellent** database schema and loading the data into the database • Use of SQL and Python together to perform complete data analysis applying **advanced** analysis methods	Project 3 (Final Project) Case Study 1 for Kaggle Competition Case Study 2 for Mini Project SQL
Write clean, modular, and efficient code following best practices and maintain a clean and logical project structure	**Coding**	**At least three of these apply:** • **Many** unused code present in the project • No functions • Naming conventions are **not** applied, making the code hard to read • **Many** hard-coded values or global variables are used • **No** consistent approach in naming, structure, and organization of files/folders	**At least four of these apply:** • **Some** unused code present in the project • Functions are either too large or perform multiple tasks • Naming conventions are **barely** applied, making the code hard to read • **Some** hard-coded values or global variables are used • **Some** files and folders organized and distributed appropriately, others need improvement for clarity	**At least four of these apply:** • **Little** unused code present in the project • Functions are modular and resusable and saved in .py files • Naming conventions are **well** applied • **Little** hard-coded values or global variables are used • **Most** files and folders organized and distributed appropriately	**All of these apply:** • **No** unused code present in the project • Functions are **perfectly** modular and resusable and saved in .py files • Naming conventions are **perfectly** applied • **No** hard-coded values or global variables are used, config files utilized instead • **All** files and folders organized and distributed appropriately	Project 1 (DW) Project 2 (EDA) Project 3 (Final Project) Mini Project Web Scraping and APIs Mini Project SQL Kaggle Competition
Save and track changes in the source code using Git and Github	**Coding**	**All of these apply:** • **Zero or one** commit **in total** • Commit messages don't provide useful information	**All of these apply:** • **At least two** commits made **during the project development** • Commit messages are **unclear** and **ambiguous**	**At least two of these apply:** • **Some** commits made **during the project development, but less than one commit per project day** • Commit messages are **clear** and **accurately** describe changes made • No separate branches used for development, if working in pairs or groups	**All of these apply:** • **At least one** commit made **per project day** • **Atomic** commits used to log changes with **accurate** and **precise** descriptions consistently across both repositories • Separate branches used for development, if working in pairs or groups	Project 1 (DW) Project 2 (EDA) Project 3 (Final Project) Mini Project Web Scraping and APIs Mini Project SQL Kaggle Competition
Document the project's features, configuration, and technical specifications	**Coding**	**All of these apply:** • No attempt made to document project specifications in a README file • No attempt made to code or fuctions	**At least two of these apply:** • **Partially** completed README provided • Functions **partially** documented with incomplete docstrings • **Few** comments provided to explain rationale, logic, or main ideas behind the code	**All of these apply:** • **Well-structured** and **clear** README file • Functions documented with **accurate** docstrings • **Enough** comments provided to explain rationale, logic, or main ideas behind the code	**All of these apply:** • **Fully comprehensive** and **well-structured** and clear README file • Functions **perfectly** documented with complete docstrings • **Clear** and **concise** comments explaining purpose and functionality of the code	Project 1 (DW) Project 2 (EDA) Project 3 (Final Project) Mini Project Web Scraping and APIs Mini Project SQL Kaggle Competition

Criteria (learning outcome)	Domain	0 - Incomplete	1- Fair	2- Good	3- Excellent	Assessment event in which applies
Build a presentation and perform a demo to deliver your final results, or the results of your group	**Data visualization and communication**	**At least two of these apply:** • Presentation **lacks a clear** structure and purpose, making it **difficult** to understand the final results • Demo is **poorly** executed and does **not effectively** communicate the final results • **Absence** of storytelling, making the presentation less engaging and memorable	**At least two of these apply:** • Presentation has **some** structure, but could benefit from **better** organization and **clearer** presentation of findings and analysis to support conclusions • Demo **adequately** communicates the final results, but could be more polished and engaging, with **improved pacing** • Storytelling is **present**, but needs to be **more effectively** used to capture the audience's attention and communicate the final results	**At least three of these apply:** • Presentation has a **clear** structure and purpose that **effectively** communicates the final results • Demo is **engaging**, **well-rehearsed**, and **effectively** communicates the final results within the allocated time frame. • Storytelling techniques are **well-incorporated** and **enhance** the audience's engagement with the presentation • Conclusions and next steps/next opportunities are included to continue improving the project	**At least four of these apply:** • Presentation is **highly compelling**, with a **clear** message and **well-structured** • Demo is **flawlessly** executed, showcasing the final results in an **engaging** and **memorable** way, and **proficiently** uses allocated time. • Storytelling techniques are **expertly** used to build a **compelling** narrative, **enhance** the audience's understanding and **emotional** connection to the final results, leaving a **lasting** impact. • Conclusions and next steps/next opportunities are included, along with strengths and limitations of the data and recommendations for further analysis. • Data visualization techniques are used in an **innovative** and **meaningful** way to support the findings and analysis. .	Project 1 (DW) Project 2 (EDA) Project 3 (Final Project) Mini Project Web Scraping and APIs Mini Project SQL Kaggle Competition

Observation: Flexibility in Applying Learning Outcomes and Criteria Points
In this agnostic project rubric, it is important to acknowledge that each project may have unique characteristics and requirements. As such, it is understood that not all learning outcomes or criteria points listed in the rubric will be applicable to every project. Students should focus on relevant outcomes and criteria for self-assessment, while teachers should consider project-specific aspects for evaluation. This approach ensures a tailored assessment aligned with the unique requirements of each project.