



WELCOME TO THE
NAAN MUDHALVAN PROJECT

PENGUIN CLASSIFICATION ANALYSIS

Team ID: NM2023TMID19767

Team Size : 5

TEAM DETAILS

Team Leader : GOKULA KANNAN V

Team member : HARIHARAN B

Team member : GOKULAKRISHNAN S

Team member : GURUMOORTHY K

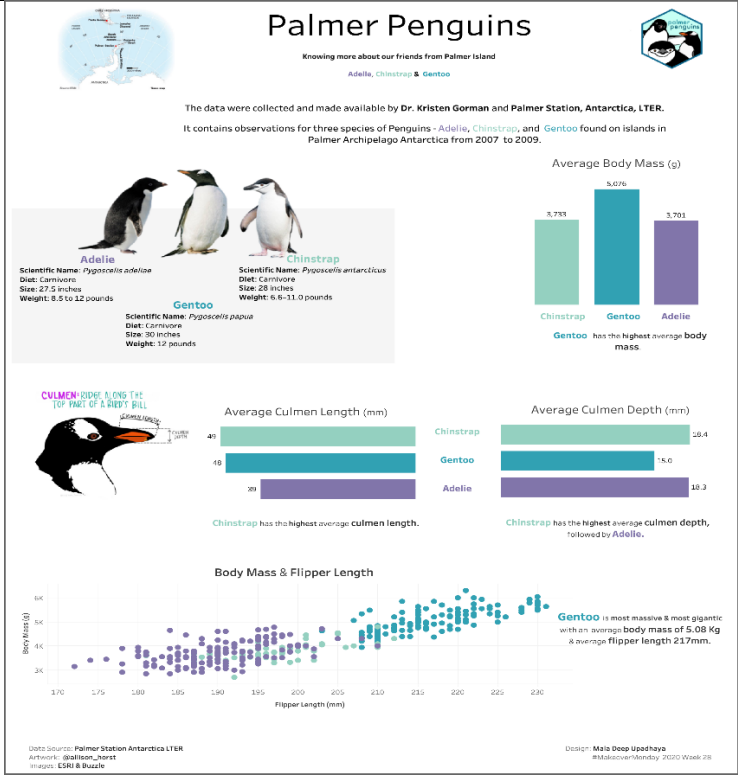
Team member : KALEESWARAN G

Project Development Phase Performance Test

Date	13 May 2023
Team ID	NM2023TMID19767
Project Name	PENGUIN CLASSIFICATION ANALYSIS

Model Performance Testing:

Project team shall fill the following information in the performance testing template.

S.No.	Parameter	Screenshot / Values
1.	Dashboard design	 <p>The screenshot displays a dashboard titled "Palmer Penguins" with the subtitle "Knowing more about our friends from Palmer Island". It lists the species: Adelle, Chinstrap, & Gentoo. The data was collected by Dr. Kristen Gorman and Palmer Station, Antarctica, LTER, from 2007 to 2009. The dashboard includes several charts and informational text:</p> <ul style="list-style-type: none"> Species Information: <ul style="list-style-type: none"> Adelle: Scientific Name: <i>Pygoscelis adelleae</i>, Diet: Carnivore, Size: 27.5 inches, Weight: 8.5 to 12 pounds. Chinstrap: Scientific Name: <i>Pygoscelis antarcticus</i>, Diet: Carnivore, Size: 28 inches, Weight: 6.8-11.0 pounds. Gentoo: Scientific Name: <i>Pygoscelis papua</i>, Diet: Carnivore, Size: 30 inches, Weight: 12 pounds. Average Body Mass (g): A bar chart showing Gentoo (5,676g), Chinstrap (5,751g), and Adelle (5,701g). Gentoo has the highest average body mass. Average Culmen Length (mm): A horizontal bar chart showing Chinstrap (38.4mm), Gentoo (35.0mm), and Adelle (33.3mm). Chinstrap has the highest average culmen length. Average Culmen Depth (mm): A horizontal bar chart showing Chinstrap (28.4mm), Gentoo (25.0mm), and Adelle (23.3mm). Chinstrap has the highest average culmen depth, followed by Adelle. Body Mass & Flipper Length: A scatter plot showing the relationship between flipper length (mm) on the x-axis and body mass (g) on the y-axis. Gentoo is noted as the most massive & most gigantic with an average body mass of 5.08 Kg & average flipper length 217mm. <p>At the bottom, it states: Data Source: Palmer Station Antarctica LTER, Artwork: @allison_cherit, Images: @DRI & @paulie. Design: Mala Deep Upadhyaya, #MakeoverMonday 2020 Week 28.</p>
		<p>The rigorous study was conducted in the islands of the Palmer Archipelago, Antarctica. These data were collected from 2007 to 2009 by Dr. Kristen Gorman with the Palmer Station Long Term Ecological Research Program, part of the US Long Term Ecological Research Network.</p>

		<p>The original GitHub repo contains the source code. You may download the dataset from Kaggle. It has two datasets, each with 344 observations. The dataset we will be using is a curated subset of the raw dataset.</p>
2.	Data Responsiveness	<p>Data responsiveness in a penguin classification analysis project refers to the ability of the dashboard to handle and display data efficiently, regardless of the volume or complexity of the dataset. Here are some considerations for ensuring data responsiveness:</p> <ul style="list-style-type: none"> ➤ Efficient data retrieval ➤ Data compression and optimization ➤ Pagination and lazy loading ➤ Data caching ➤ Asynchronous processing ➤ Data aggregation and summarization ➤ Progressive loading ➤ Performance monitoring and optimization
3.	Utilization of Data Filters	<p>Data filters play a crucial role in a penguin classification analysis project as they allow users to refine and explore the data based on specific criteria. Here are some key ways to utilize data filters effectively:</p> <ul style="list-style-type: none"> ➤ Species filter ➤ Age filter ➤ Geographical filter ➤ Time filter ➤ Attribute filter ➤ Multi-select filters ➤ Interactive filter widgets ➤ Filter-dependent visualizations
4.	Effective User Story	<p>User Story: As a data analyst, I want to explore and classify penguin data accurately to gain insights into different penguin species and their characteristics, enabling better understanding and conservation efforts.</p> <p>Acceptance Criteria:</p> <ul style="list-style-type: none"> ➤ The system should provide a user-friendly interface to interact with the penguin classification analysis. ➤ The user should be able to upload and import penguin data in a standardized format.

		<ul style="list-style-type: none"> ➤ The system should preprocess the data by handling missing values and outliers appropriately. ➤ The user should be able to visualize the distribution of penguin species based on attributes such as bill length, body mass, and flipper size. ➤ The system should provide interactive filters to explore the data based on species, age, geographical location, or time period. ➤ The user should be able to apply machine learning algorithms for penguin classification, such as decision trees, logistic regression, or neural networks. ➤ The system should evaluate the classification models using appropriate metrics such as accuracy, precision, recall, and F1 score. ➤ The user should have the ability to compare and visualize the performance of different classification models.
5.	Descriptive Reports	<p>Descriptive reports for a penguin classification analysis project provide a comprehensive overview of the data, analysis methods, and key findings. These reports aim to communicate the analysis process and results to stakeholders, researchers, or decision-makers. Here are some components to include in descriptive reports for a penguin classification analysis project:</p> <ul style="list-style-type: none"> ➤ Executive Summary ➤ Introduction ➤ Data Description ➤ Methodology ➤ Results ➤ Interpretation and Insights ➤ Discussion