

UNVEILING THE STARS:

AN EXPLORATORY STUDY ON NASA ASTRONAUTS



By Ananwita Sarkar

PROJECT DESCRIPTION

The business problem at hand is to harness the vast reservoir of data on NASA astronauts and convert it into a valuable resource for researchers, space enthusiasts, and dreamers. This project aims to unveil the compelling stories and collective achievements of NASA's astronaut corps by conducting a comprehensive analysis of their demographics, careers, and accomplishments. By applying state-of-the-art data analysis tools and techniques, we seek to extract meaningful insights, uncovering patterns and trends that illuminate the diverse tapestry of these space pioneers. Ultimately, the objective is to transform raw information into actionable knowledge, facilitating a deeper understanding of the individuals who have ventured beyond our atmosphere and their profound impact on the ongoing saga of human exploration.

DATA COLLECTION

The dataset used for this analysis was obtained from Kaggle, and it consisted of [briefly describe the dataset, its purpose, and relevant details]. During the initial examination of the dataset, it was observed that there were 302 missing values in both the 'Death_Mission' and 'Death_Date' columns.

DATA CLEANING AND MISSING VALUE HANDLING

To address the issue of missing values, the following steps were taken using Python:

- **Filtering Null Values:** To address missing data, rows with null 'Death_Mission' but non-null 'Death_Date' were identified using `.index` and then removed from 'df' using `.drop()` (`inplace=True`), ensuring data consistency.
- **Handling Missing Values:** Rows with missing 'Alma_Mater' values were removed using `.dropna()` for a dataset with known education backgrounds.
- **Data Type Conversion:** The 'Birth_Date' and 'Death_Date' columns were converted to datetime with `dayfirst=True` for consistent date handling.

DATA PREPARATION AND PRE-PROCESSING

The data preparation and preprocessing phase in this project encompassed several critical steps that ensured data quality, accessibility, and integrity:

- **Exporting the DataFrame:** Following thorough DataFrame cleaning in Jupyter Notebook, it was essential to export the DataFrame to a CSV file. This step preserves processed data, ensuring future access and collaboration. Notably, setting 'index' to False omits index inclusion for cleaner data representation.
- **Data Download and Import:** After exporting, the CSV dataset was downloaded, signifying the shift from local to external data handling. It was then imported into phpMyAdmin for database integration and analysis.
- **Database Connection:** The database connection involved creating a MySQL table, uploading datasets, renaming the table, loading the SQL extension in Jupyter Notebook, and specifying the MySQL connection string for real-time data access.

RESEARCH QUESTIONS

- Counting Astronauts by Status
- Counting Astronauts by Military Branch.
- Counting Astronauts by Military Branch.
- Top 5 Military Ranks Among Astronauts.
- Count of Astronauts by Gender.
- Average Life Expectancy of Astronauts.
- Average Life Expectancy of Female Astronauts.
- Average Life Expectancy of Male Astronauts.
- Top 10 Graduate Majors Among Astronauts.
- Astronaut Education Statistics.
- Top 5 States of Birth for Astronauts.
- Average Space Flights and Spacewalks per Astronaut.

EXPLORATORY DATA ANALYSIS(EDA)

TASK 1: RETRIEVING ALL DATA

Retrieved dataset from a reliable source, providing the basis for the project. The dataset contains comprehensive information about NASA astronauts, enabling in-depth analysis of their demographics, careers, and achievements.

TASK 2: COUNTING ASTRONAUTS BY STATUS

Status	Number
Deceased	48
Retired	216
Active	50
Management	36

Insights: The analysis of astronaut status reveals that 48 astronauts are deceased, highlighting their significant contributions and sacrifices. A majority of 216 astronauts have retired, while 50 remain active in missions, emphasizing the ongoing nature of space exploration. Additionally, 36 astronauts hold management roles, indicating their leadership and administrative responsibilities within NASA's astronaut program.

TASK 3: COUNTING ASTRONAUTS BY MILITARY BRANCH.

Military_Branch	Number
US Navy (Retired)	100
US Navy	37
US Air Force (Retired)	108
US Air Force	26
US Marine Corps (Retired)	29
US Army (Retired)	21
US Coast Guard (Retired)	4
US Army	10
US Marine Corps	3
US Air Force Reserves	3
US Naval Reserves (Retir...	2
US Naval Reserves	2
US Air Force Reserves (R...	3
US Marine Corps Reserves	2

Insights: The analysis of astronauts by military branch reveals a notable presence of retired military personnel, with the US Navy (Retired) and US Air Force (Retired) being the most common. Active military branches like the US Navy and US Air Force are also represented. Additionally, a smaller number of astronauts have backgrounds in reserve military branches, including the US Air Force Reserves, US Naval Reserves, US Marine Corps Reserves, and US Air Force Reserves (Retired), showcasing the diversity of military experiences within the astronaut corps.

TASK 4: TOP 5 MILITARY RANKS AMONG ASTRONAUTS.

Military_Rank	Number
Colonel	155
Captain	115
Commander	26
Lieutenant Colonel	17
Major General	12

Insights: It reveals that Colonels and Captains are the most prevalent, with 155 and 115 astronauts, respectively which indicates the significance of these ranks in the astronaut corps.

Commanders and Lieutenant Colonels follow with 26 and 17 astronauts which indicates that these ranks are vital for leadership roles and responsibilities within space missions.

Major Generals are also represented, with 12 astronauts in this high-ranking role which suggests that some astronauts attain senior military positions before becoming space pioneers.

TASK 5: COUNT OF ASTRONAUTS BY GENDER.

Gender	Number
Male	301
Female	49

Insights: The analysis reveals the gender distribution among astronauts is 301 males and 49 females. This data illustrates the historical gender disparity in the astronaut corps, with males in the majority. However, the presence of 49 female astronauts reflects positive progress towards greater gender diversity and inclusivity in space exploration.

TASK 6: AVERAGE LIFE EXPECTANCY OF ASTRONAUTS.

Average_Life_Expectancy
69

Insights: This analysis calculates that the average life expectancy of astronauts is 69 years. This data provides an overview of the expected lifespan of astronauts and contributes to discussions about the impact of space missions on astronauts' longevity.

TASK 7: AVERAGE LIFE EXPECTANCY OF FEMALE ASTRONAUTS.

Female_Average_Life_Expectancy

62

Insights: The analysis calculates the average life expectancy of female astronauts is 62 years. This data sheds light on the expected lifespan of female astronauts and may be of interest in discussions regarding the unique challenges and experiences faced by women in space exploration.

TASK 8: AVERAGE LIFE EXPECTANCY OF MALE ASTRONAUTS.

Male_Average_Life_Expectancy
70

Insights: The analysis reveals that the average life expectancy for male astronauts is 70 years. This data provides specific information about the expected lifespan of male astronauts and contributes to discussions about the effects of space missions on their longevity.

TASK 9: TOP 10 GRADUATE MAJORS AMONG ASTRONAUTS.

Graduate_Major	Number
Aeronautical Engineering	34
Aerospace Engineering	26
Physics	17
Medicine	16
Mechanical Engineering	14
Electrical Engineering	9
Astronautics	8
Aeronautics & Astronautics	7
Aeronautics	7
Aviation Systems	7

Insights: The analysis of the top 10 graduate majors among astronauts reveals a prevalence of engineering backgrounds with Aeronautical Engineering and Aerospace Engineering as the most common, represented by 34 and 26 astronauts, respectively. The list also includes diverse scientific disciplines, with Physics and Medicine prominent. Technical fields like Mechanical and Electrical Engineering are noteworthy. Additionally, some astronauts have specialized degrees in astronautics or space-related fields, illustrating their specific knowledge in space exploration. These insights highlight the wide array of educational backgrounds within the astronaut corps.

TASK 10: ASTRONAUT EDUCATION STATISTICS.

Number_of_Astronauts	Astronauts_with_Undergraduate_Degrees	Astronauts_with_Graduate_Degrees
350	350	350

Insights: The analysis reveals that out of a total of 350 astronauts, each astronaut holds both an undergraduate and a graduate degree. This data highlights the stringent educational qualifications required for individuals within the astronaut corps, with every astronaut meeting these educational standards.

TASK 11: TOP 5 STATES OF BIRTH FOR ASTRONAUTS.

state	astro_number
NY	30
CA	25
TX	23
OH	21
PA	19

Insights: The analysis identifies the top 5 states of birth for astronauts. New York leads with 30 astronauts, followed by California with 25, Texas with 23, Ohio with 21, and Pennsylvania with 19. These insights showcase the states that have produced a significant number of astronauts.

TASK 12: AVERAGE SPACE FLIGHTS AND SPACEWALKS PER ASTRONAUT.

Average_Number_Of_Space_Flight	Average_Number_Of_Space_Walks
2.40	1.27

Insights: The analysis calculates the average number of space flights and spacewalks per astronaut. On average, each astronaut has participated in around 2.40 space flights, emphasizing the frequency of their involvement in space missions. Additionally, astronauts have an average of approximately 1.27 spacewalks, highlighting their experiences in extravehicular activities. These insights offer a glimpse into astronauts' career involvement in space exploration.

SUGGESTIONS

- Maintain the source of the dataset and ensure it is regularly updated to reflect the most current information.
- Further investigate the reasons behind different status categories and their implications for astronaut careers.
- Explore how military experience influences astronaut selection and performance in space missions.
- Analyze how different military ranks contribute to leadership and responsibilities in space missions.
- Examine strategies to increase female representation and enhance gender diversity in future astronaut selections.
- Investigate the factors that contribute to astronauts' life expectancies and explore potential ways to improve their well-being.
- Investigate the factors that contribute to astronauts' life expectancies and explore potential ways to improve their well-being.
- Conduct further research into the unique challenges faced by female astronauts and their impact on life expectancy.
- Examine the factors contributing to longer life expectancies among male astronauts and their potential impact on space missions.
- Investigate how various educational backgrounds contribute to success in space missions and explore the importance of diverse expertise.
- Continue to maintain high educational standards while ensuring diversity in educational backgrounds.
- Analyze why these states have produced a significant number of astronauts and explore potential outreach programs in other regions.
- Investigate the impact of frequent space missions and spacewalks on astronauts' careers and well-being.

THANK YOU