

```
In [1]: import pandas as pd
```

Pandas allows us to analyze big data and make conclusions based on statistical theories. Pandas can clean messy data sets, and make them readable and relevant. Relevant data is very important in data science.

```
In [3]: df = pd.read_csv(r'C:\Users\hp\Downloads\nasa.csv')
```

Steps we are following to clean data that will be used for data analysis using SQL:

1. Above we uploaded the CSV data 'nasa.csv' into data frame df
2. Understand / Inspect the data
3. Check for null values
4. Using sum() function to check sum of all null values
5. Droping null values
6. Creating a new CSV file to be uploaded in SQL database

```
In [4]: df
```

Out[4]:

	Name	Year	GroupNum	Status	Birth_Date	Birth_Place	Gender	Alma_Mater	Under
0	Alan B. Shepard Jr.	1959	1	Deceased	18-11-1923	East Derry, NH	Male	US Naval Academy	
1	Alan G. Poindexter	1998	17	Deceased	05-11-1961	Pasadena, CA	Male	Georgia Institute of Technology; US Naval Post...	Aeros...
2	Alan L. Bean	1963	3	Deceased	15-03-1932	Wheeler, TX	Male	University of Texas	
3	Albert Sacco Jr.	1963	3	Retired	03-05-1949	Boston, MA	Male	Northeastern University; MIT	Cher...
4	Alfred M. Worden	1966	5	Retired	07-02-1932	Jackson, MI	Male	US Military Academy; University of Michigan	
...
352	William M. Shepherd	1984	10	Retired	26-07-1949	Oak Ridge, TN	Male	US Naval Academy; MIT	Aeros...
353	William R. Pogue	1966	5	Retired	23-01-1930	Okemah, OK	Male	Oklahoma Baptist University; Oklahoma State Un...	
354	William S. McArthur Jr.	1990	3	Management	26-07-1951	Laurinburg, NC	Male	US Military Academy; Georgia Institute of Tech...	A
355	Winston E. Scott	1992	14	Retired	06-08-1950	Miami, FL	Male	Florida State University; US Naval Postgraduat...	
356	Yvonne D. Cagle	1996	16	Management	24-04-1959	West Point, NY	Female	San Francisco State University	

357 rows × 19 columns



In [5]: df.describe()

Out[5]:

	Year	GroupNum	Space_Flights	Space_Flight_hr	Space_Walks	Space_Walks_hr
count	357.000000	357.000000	357.000000	357.000000	357.000000	357.000000
mean	1984.739496	11.249300	2.364146	1249.266106	1.246499	7.707283
std	13.347850	5.221179	1.428700	1896.759857	2.056989	13.367973
min	1959.000000	1.000000	0.000000	0.000000	0.000000	0.000000
25%	1978.000000	8.000000	1.000000	289.000000	0.000000	0.000000
50%	1987.000000	12.000000	2.000000	590.000000	0.000000	0.000000
75%	1996.000000	16.000000	3.000000	1045.000000	2.000000	12.000000
max	2009.000000	20.000000	7.000000	12818.000000	10.000000	67.000000

In [6]:

df.head()

Out[6]:

	Name	Year	GroupNum	Status	Birth_Date	Birth_Place	Gender	Alma_Mater	Undergradu...
0	Alan B. Shepard Jr.	1959	1	Deceased	18-11-1923	East Derry, NH	Male	US Naval Academy	Nava...
1	Alan G. Poindexter	1998	17	Deceased	05-11-1961	Pasadena, CA	Male	Georgia Institute of Technology; US Naval Post...	Aerospace E...
2	Alan L. Bean	1963	3	Deceased	15-03-1932	Wheeler, TX	Male	University of Texas	Ae... E...
3	Albert Sacco Jr.	1963	3	Retired	03-05-1949	Boston, MA	Male	Northeastern University; MIT	Chemical E...
4	Alfred M. Worden	1966	5	Retired	07-02-1932	Jackson, MI	Male	US Military Academy; University of Michigan	Milita...

In [7]:

df.isnull()

Out[7]:

	Name	Year	GroupNum	Status	Birth_Date	Birth_Place	Gender	Alma_Mater	Undergraduate_Major
0	False	False		False	False	False	False	False	False
1	False	False		False	False	False	False	False	False
2	False	False		False	False	False	False	False	False
3	False	False		False	False	False	False	False	False
4	False	False		False	False	False	False	False	False
...
352	False	False		False	False	False	False	False	False
353	False	False		False	False	False	False	False	False
354	False	False		False	False	False	False	False	False
355	False	False		False	False	False	False	False	False
356	False	False		False	False	False	False	False	False

357 rows × 19 columns

In [8]: `df.isnull().sum()`

Out[8]:

Name	0
Year	0
GroupNum	0
Status	0
Birth_Date	0
Birth_Place	0
Gender	0
Alma_Mater	1
Undergraduate_Major	0
Graduate_Major	0
Military_Rank	0
Military_Branch	0
Space_Flights	0
Space_Flight_hr	0
Space_Walks	0
Space_Walks_hr	0
Missions	0
Death_Date	303
Death_Mission	309
dtype: int64	

Data Cleaning and Filtering Operations.

1. Remove rows with missing values in the 'Alma_Mater' column using `dropna()` and applies the changes directly to the DataFrame 'df' with `inplace=True`.
2. Filter rows where 'Death_Mission' column is null and 'Death_Date' column is not null using the `.index` method. Then, remove those rows from the DataFrame 'df' by using the `.drop()` function with the `inplace=True` parameter.

3. Use pd.to_datetime() with the dayfirst=True parameter to convert the 'Birth_Date' and 'Death_Date' columns to datetime format with day-first parsing separately.

```
In [12]: df.dropna(subset=['Alma_Mater'], inplace=True)
rows_to_drop = df[df['Death_Mission'].isnull() & df['Death_Date'].notnull()].index
```

```
In [13]: rows_to_drop
```

```
Out[13]: Int64Index([], dtype='int64')
```

```
In [14]: df.drop(rows_to_drop, inplace=True)
df['Birth_Date'] = pd.to_datetime(df['Birth_Date'], dayfirst=True)
df['Death_Date'] = pd.to_datetime(df['Death_Date'], dayfirst=True)
```

Now creating a new CSV file which will be SQL database

```
In [18]: df.to_csv('astronauts.csv', index=False)
```

```
In [22]: import sqlite3
```

```
In [23]: astronauts = pd.read_csv('astronauts.csv')
```

```
In [24]: astronauts
```

Out[24]:

	Name	Year	GroupNum	Status	Birth_Date	Birth_Place	Gender	Alma_Mater	Under
0	Alan B. Shepard Jr.	1959	1	Deceased	1923-11-18	East Derry, NH	Male	US Naval Academy	
1	Alan G. Poindexter	1998	17	Deceased	1961-11-05	Pasadena, CA	Male	Georgia Institute of Technology; US Naval Post...	Aeros...
2	Alan L. Bean	1963	3	Deceased	1932-03-15	Wheeler, TX	Male	University of Texas	
3	Albert Sacco Jr.	1963	3	Retired	1949-05-03	Boston, MA	Male	Northeastern University; MIT	Cher...
4	Alfred M. Worden	1966	5	Retired	1932-02-07	Jackson, MI	Male	US Military Academy; University of Michigan	
...
345	William M. Shepherd	1984	10	Retired	1949-07-26	Oak Ridge, TN	Male	US Naval Academy; MIT	Aeros...
346	William R. Pogue	1966	5	Retired	1930-01-23	Okemah, OK	Male	Oklahoma Baptist University; Oklahoma State Un...	
347	William S. McArthur Jr.	1990	3	Management	1951-07-26	Laurinburg, NC	Male	US Military Academy; Georgia Institute of Tech...	A
348	Winston E. Scott	1992	14	Retired	1950-08-06	Miami, FL	Male	Florida State University; US Naval Postgraduat...	
349	Yvonne D. Cagle	1996	16	Management	1959-04-24	West Point, NY	Female	San Francisco State University	

350 rows × 19 columns

In [25]: `conn = sqlite3.connect('my_database.sqlite')`In [26]: `astronauts.to_sql('astronauts', conn)`

Out[26]: 350

Checking if ipython_sql library is present to run SQL queries directly on Jupyter Notebook

In [31]: pip install ipython-sql

```
Requirement already satisfied: ipython-sql in c:\users\hp\anaconda3\lib\site-packages (0.5.0)
Requirement already satisfied: prettytable in c:\users\hp\anaconda3\lib\site-packages (from ipython-sql) (3.9.0)
Requirement already satisfied: ipython in c:\users\hp\anaconda3\lib\site-packages (from ipython-sql) (8.12.0)
Requirement already satisfied: sqlalchemy>=2.0 in c:\users\hp\anaconda3\lib\site-packages (from ipython-sql) (2.0.22)
Requirement already satisfied: sqlparse in c:\users\hp\anaconda3\lib\site-packages (from ipython-sql) (0.4.4)
Requirement already satisfied: six in c:\users\hp\anaconda3\lib\site-packages (from ipython-sql) (1.16.0)
Requirement already satisfied: ipython-genutils in c:\users\hp\anaconda3\lib\site-packages (from ipython-sql) (0.2.0)
Requirement already satisfied: typing-extensions>=4.2.0 in c:\users\hp\anaconda3\lib\site-packages (from sqlalchemy>=2.0->ipython-sql) (4.6.3)
Requirement already satisfied: greenlet!=0.4.17 in c:\users\hp\anaconda3\lib\site-packages (from sqlalchemy>=2.0->ipython-sql) (2.0.1)
Requirement already satisfied: backcall in c:\users\hp\anaconda3\lib\site-packages (from ipython->ipython-sql) (0.2.0)
Requirement already satisfied: decorator in c:\users\hp\anaconda3\lib\site-packages (from ipython->ipython-sql) (5.1.1)
Requirement already satisfied: jedi>=0.16 in c:\users\hp\anaconda3\lib\site-packages (from ipython->ipython-sql) (0.18.1)
Requirement already satisfied: matplotlib-inline in c:\users\hp\anaconda3\lib\site-packages (from ipython->ipython-sql) (0.1.6)
Requirement already satisfied: pickleshare in c:\users\hp\anaconda3\lib\site-packages (from ipython->ipython-sql) (0.7.5)
Requirement already satisfied: prompt-toolkit!=3.0.37,<3.1.0,>=3.0.30 in c:\users\hp\anaconda3\lib\site-packages (from ipython->ipython-sql) (3.0.36)
Requirement already satisfied: pygments>=2.4.0 in c:\users\hp\anaconda3\lib\site-packages (from ipython->ipython-sql) (2.15.1)
Requirement already satisfied: stack-data in c:\users\hp\anaconda3\lib\site-packages (from ipython->ipython-sql) (0.2.0)
Requirement already satisfied: traitlets>=5 in c:\users\hp\anaconda3\lib\site-packages (from ipython->ipython-sql) (5.7.1)
Requirement already satisfied: colorama in c:\users\hp\anaconda3\lib\site-packages (from ipython->ipython-sql) (0.4.6)
Requirement already satisfied: wcwidth in c:\users\hp\anaconda3\lib\site-packages (from prettytable->ipython-sql) (0.2.5)
Requirement already satisfied: parso<0.9.0,>=0.8.0 in c:\users\hp\anaconda3\lib\site-packages (from jedi>=0.16->ipython->ipython-sql) (0.8.3)
Requirement already satisfied: executing in c:\users\hp\anaconda3\lib\site-packages (from stack-data->ipython->ipython-sql) (0.8.3)
Requirement already satisfied: asttokens in c:\users\hp\anaconda3\lib\site-packages (from stack-data->ipython->ipython-sql) (2.0.5)
Requirement already satisfied: pure-eval in c:\users\hp\anaconda3\lib\site-packages (from stack-data->ipython->ipython-sql) (0.2.2)
Note: you may need to restart the kernel to use updated packages.
```

Loading the library

In [32]: `%load_ext sql`

Now, we will do analysis using SQL:

1. We will connect our database which contains our CSV file, '%sql' magic command is used to run SQL queries on notebook
2. Run SQL queries to do analysis on our data (detailed explanation is given below each code block).

```
In [33]: %sql sqlite:///my_database.sqlite  
%sql SELECT * FROM astronauts;
```

```
* sqlite:///my_database.sqlite  
Done.
```

index	Name	Year	GroupNum	Status	Birth Date	Birth Place	Gender	Alma Mater
0	Alan B. Shepard Jr.	1959	1	Deceased	1923-11-18	East Derry, NH	Male	US Naval Academy
1	Alan G. Poindexter	1998	17	Deceased	1961-11-05	Pasadena, CA	Male	Georgia Institute of Technology US Naval Postgraduate School
2	Alan L. Bean	1963	3	Deceased	1932-03-15	Wheeler, TX	Male	University of Texas
3	Albert Sacco Jr.	1963	3	Retired	1949-05-03	Boston, MA	Male	Northeastern University M
4	Alfred M. Worden	1966	5	Retired	1932-02-07	Jackson, MI	Male	US Military Academy University of Michigan
5	Alvin B. Drew Jr.	2000	18	Active	1962-11-05	Washington, DC	Male	US Air Force Academy Embry-Riddle Aeronautical University
6	Andrew J. Feustel	2000	18	Active	1965-08-25	Lancaster, PA	Male	Purdue University Queen's University Canada
7	Andrew M. Allen	1987	12	Retired	1955-08-04	Philadelphia, PA	Male	Villanova University University of Florida
8	Andrew S. W. Thomas	1992	14	Management	1951-12-18	Adelaide, Australia	Male	University of Adelaide
9	Anna L. Fisher	1978	8	Management	1949-08-24	New York, NY	Female	University of California-Los Angeles
10	Anthony W. England	1967	6	Retired	1942-05-15	Indianapolis, IN	Male	M
11	Barbara R. Morgan	1998	17	Retired	1951-11-28	Fresno, CA	Female	Stanford University
12	Barry E. Wilmore	2000	18	Active	1962-12-29	Murfreesboro, TN	Male	Tennessee Technological University University of Tennessee

index	Name	Year	GroupNum	Status	Birth_Date	Birth_Place	Gender	Alma_Mate
13	Bernard A. Harris Jr.	1990	13	Retired	1956-06-26	Temple, TX	Male	University of Houston, Texas Tech University
14	Bill Nelson	1990	13	Retired	1942-09-29	Miami, FL	Male	Yale University, University of Virgin
15	Bonnie J. Dunbar	1980	9	Retired	1949-03-03	Sunnyside, WA	Female	University of Washington, University of Houston
16	Brent W. Jett	1992	14	Retired	1958-10-05	Pontiac, MI	Male	US Naval Academy; US Naval Postgraduate School
17	Brewster H. Shaw Jr.	1978	8	Retired	1945-05-16	Cass City, MI	Male	University of Wisconsin
18	Brian Duffy	1985	11	Retired	1953-06-20	Boston, MA	Male	US Air Force Academy, University of Southern California
19	Brian T. O'Leary	1967	6	Deceased	1940-01-27	Boston, MA	Male	William College, Georgetown University, University of California Berkeley
20	Bruce E. Melnick	1987	12	Retired	1949-12-05	New York, NY	Male	US Coast Guard Academy, West Florida University
21	Bruce McCandless II	1966	5	Retired	1937-06-08	Boston, MA	Male	US Naval Academy, Stanford University, University of Houston, Clear Lake
22	Bryan D. O'Connor	1980	9	Retired	1946-09-06	Orange, CA	Male	US Naval Academy, West Florida University

index	Name	Year	GroupNum	Status	Birth Date	Birth Place	Gender	Alma Mater
23	Buzz Aldrin	1963	3	Retired	1930-01-20	Montclair, NJ	Male	US Military Academy; M
24	Byron K. Lichtenberg	1963	3	Retired	1948-02-19	Stroudsburg, PA	Male	Brown University; M
25	C. Gordon Fullerton	1969	7	Retired	1936-10-11	Rochester, NY	Male	California Institute of Technology
26	C. Michael Foale	1987	12	Active	1957-01-06	Louth, England	Male	Cambridge University
27	Carl E. Walz	1990	13	Retired	1955-09-06	Cleveland, OH	Male	Kent State University; John Carroll University
28	Carl J. Meade	1985	11	Retired	1950-11-16	Chanute Air Force Base, IL	Male	University of Texas at San Antonio; California Institute of Technology
29	Carlos I. Noriega	1995	15	Retired	1959-10-08	Lima, Peru	Male	University of Southern California; US Naval Postgraduate School
30	Catherine G. Coleman	1992	14	Active	1960-12-14	Charleston, SC	Female	Michigan University; Massachusetts
31	Charles A. Bassett II	1963	3	Deceased	1931-12-30	Dayton, OH	Male	Texas Technological College
32	Charles Conrad Jr.	1962	2	Deceased	1930-05-02	Philadelphia, PA	Male	Princeton University
33	Charles D. Gemar	1985	11	Retired	1955-08-04	Yankton, SD	Male	US Military Academy
34	Charles D. Walker	1985	11	Retired	1948-08-29	Bedford, IN	Male	Purdue University
35	Charles E. Brady Jr.	1992	14	Deceased	1951-08-12	Pinehurst, NC	Male	University of North Carolina at Chapel Hill

index	Name	Year	GroupNum	Status	Birth Date	Birth Place	Gender	Alma Mater
36	Charles F. Bolden Jr.	1980	9	Management	1946-08-19	Columbia, SC	Male	Duke University
37	Charles J. Camarda	1996	16	Management	1952-05-08	Queens, NY	Male	Polytechnic Institute of Brooklyn, Georgia, Washington University, Virgin Polytechnic Institute
38	Charles J. Precourt	1990	13	Retired	1955-06-29	Waltham, MA	Male	US Air Force Academy; Golden Gate University; US Naval War College
39	Charles Lacy Veach	1984	10	Deceased	1944-09-18	Chicago, IL	Male	US Air Force Academy
40	Charles M. Duke Jr.	1966	5	Retired	1935-10-03	Charlotte, NC	Male	US Naval Academy; M
41	Charles O. Hobaugh	1996	16	Retired	1961-11-05	Bar Harbor, ME	Male	US Naval Academy
42	Christopher J. Cassidy	2004	19	Active	1970-01-04	Salem, MA	Male	US Naval Academy; M
43	Christopher J. Ferguson	1998	17	Retired	1961-09-01	Philadelphia, PA	Male	Drexel University; US Naval Postgraduate School
44	Christopher J. Loria	1996	16	Retired	1960-07-09	Belmont, MA	Male	US Naval Academy; Harvard University
45	Clayton C. Anderson	1998	17	Retired	1959-02-23	Omaha, NE	Male	Hastings College; Iowa State University
46	Clifton C. Williams Jr.	1963	3	Deceased	1932-09-26	Mobile, AL	Male	Auburn University
47	Curtis L. Brown Jr.	1987	12	Retired	1956-03-11	Elizabethtown, NC	Male	US Air Force Academy

index	Name	Year	GroupNum	Status	Birth_Date	Birth_Place	Gender	Alma_Mat
48	Dale A. Gardner	1978	8	Retired	1948-11-08	Fairmont, MN	Male	University of Illinois
49	Daniel C. Brandenstein	1978	8	Retired	1943-01-17	Watertown, WI	Male	University of Wisconsin
50	Daniel C. Burbank	1996	16	Active	1961-07-27	Machester, CT	Male	US Coast Guard Academy; Embry-Riddle Aeronautical University
51	Daniel M. Tani	1996	16	Retired	1961-02-01	Ridley Park, PA	Male	MIT
52	Daniel T. Barry	1992	14	Retired	1953-12-30	Norwalk, CT	Male	Cornell University; Princeton University; University of Miami
53	Daniel W. Bursch	1990	13	Retired	1957-07-25	Bristol, PA	Male	US Naval Academy; US Naval Postgraduate School
54	David A. Wolf	1990	13	Retired	1956-08-23	Indianapolis, IN	Male	Purdue University; Indiana University
55	David C. Hilmers	1980	9	Retired	1950-01-28	Clinton, IA	Male	Cornell University; US Naval Postgraduate School
56	David C. Leestma	1980	9	Management	1949-05-06	Muskegon, MI	Male	US Naval Academy; US Naval Postgraduate School
57	David M. Brown	1996	16	Deceased	1956-04-16	Arlington, VA	Male	College of William and Mary; Eastern Virginia Medical School

index	Name	Year	GroupNum	Status	Birth_Date	Birth_Place	Gender	Alma_Mat
58	David M. Walker	1978	8	Deceased	1944-05-20	Columbus, GA	Male	US Naval Academy
59	David R. Scott	1963	3	Retired	1932-06-06	San Antonio, TX	Male	US Military Academy; Massachusetts Institute of Technology
60	Dominic A. Antonelli	2000	18	Active	1967-08-23	Detroit, MI	Male	University of Washington
61	Dominic L. Gorie	1995	15	Retired	1957-05-02	Lake Charles, LA	Male	US Naval Academy; University of Tennessee
62	Don L. Lind	1966	5	Retired	1930-05-18	Midvale, UT	Male	University of Utah; University of California, Berkeley
63	Donald A. Thomas	1990	13	Retired	1955-05-06	Cleveland, OH	Male	Case Western Reserve University; Cornell University
64	Donald E. Williams	1978	8	Retired	1958-02-13	Lafayette, IN	Male	Purdue University
65	Donald H. Peterson	1969	7	Retired	1933-10-22	Winona, MS	Male	US Military Academy; University of Air Force Institute of Technology
66	Donald K. Slayton	1959	1	Deceased	1924-03-01	Sparta, WI	Male	University of Minnesota
67	Donald L. Holmquest	1967	6	Retired	1939-04-07	Dallas, TX	Male	Southern Methodist University; Baylor University; University of Houston
68	Donald R. McMonagle	1987	12	Retired	1952-05-14	Flint, MI	Male	US Air Force Academy; California State University, Fresno
69	Donald R. Pettit	1996	16	Active	1955-04-20	Silverton, OR	Male	Oregon State University; University of Arizona

index	Name	Year	GroupNum	Status	Birth_Date	Birth_Place	Gender	Alma_Mat
70	Donn F. Eisele	1963	3	Deceased	1930-05-23	Columbus, OH	Male	US Naval Academy; US Air Force Institute of Technology
71	Dorothy M. Metcalf-Lindenberger	2004	19	Active	1975-05-02	Colorado Springs, CO	Female	Whitman College
72	Douglas G. Hurley	2000	18	Active	1966-10-21	Endicott, NY	Male	Tulane University
73	Douglas H. Wheelock	1998	17	Active	1960-05-05	Binghamton, NY	Male	US Military Academy; Georgia Institute of Technology
74	Duane E. Graveline	1965	4	Retired	1931-03-02	Newport, VT	Male	University of Vermont; Johns Hopkins University
75	Duane G. Carey	1996	16	Retired	1957-04-30	St. Paul, MN	Male	University of Minnesota; Minneapolis
76	E. Michael Fincke	1996	16	Active	1967-03-14	Pittsburgh, PA	Male	MIT; Stanford University; University of Houston; Clear Lake
77	Edgar D. Mitchell	1966	5	Retired	1930-09-17	Hereford, TX	Male	Carnegie Mellon University; US Naval Postgraduate School; M
78	Edward G. Gibson	1965	4	Retired	1936-11-08	Buffalo, NY	Male	University of Rochester; California Institute of Technology
79	Edward G. Givens Jr.	1966	5	Deceased	1930-01-05	Quanah, TX	Male	US Naval Academy
80	Edward H. White II	1962	2	Deceased	1930-11-14	San Antonio, TX	Male	US Military Academy; University of Michigan
81	Edward T. Lu	1995	15	Retired	1963-07-01	Springfield, MA	Male	Cornell University; Stanford University
82	Eileen M. Collins	1990	13	Retired	1959-11-19	Elmira, NY	Female	Syracuse University

index	Name	Year	GroupNum	Status	Birth_Date	Birth_Place	Gender	Alma_Mate
83	Ellen Ochoa	1990	13	Management	1958-05-10	Los Angeles, CA	Female	Stanford University Webster University
84	Ellen S. Baker	1984	10	Retired	1953-04-27	Fayetteville, NC	Female	Stanford University New York University Buffalo Cornell University University of Texas
85	Elliot M. See Jr.	1962	2	Deceased	1927-07-23	Dallas, TX	Male	US Merchant Marine Academy University of California-Los Angeles
86	Ellison S. Onizuka	1978	8	Deceased	1946-06-24	Kealakekua, HI	Male	University of Colorado
87	Eric A. Boe	2000	18	Active	1964-10-01	Miami, FL	Male	US Air Force Academy Georgia Institute of Technology
88	Eugene A. Cernan	1963	3	Deceased	1934-03-14	Chicago, IL	Male	Purdue University; US Naval Postgraduate School
89	F. Andrew Gaffney	1963	3	Retired	1946-06-09	Carlsbad, NM	Male	University of California-Berkeley University of New Mexico
90	F. Curtis Michel	1965	4	Retired	1934-06-05	LaCrosse, WI	Male	California Institute of Technology
91	Francis R. Scobee	1978	8	Deceased	1939-05-19	Cle Elum, WA	Male	University of Arizona
92	Frank Borman	1962	2	Retired	1928-03-14	Gary, IN	Male	US Military Academy California Institute of Technology

index	Name	Year	GroupNum	Status	Birth_Date	Birth_Place	Gender	Alma_Mat
93	Frank L. Culbertson Jr.	1984	10	Retired	1949-05-15	Charleston, SC	Male	US Naval Academy
94	Franklin R. Chang-Diaz	1980	9	Retired	1950-04-05	San Jose, Costa Rica	Male	University of Connecticut M
95	Fred W. Haise Jr.	1966	5	Retired	1933-11-14	Biloxi, MS	Male	University of Oklahoma
96	Fred W. Leslie	1966	5	Retired	1951-12-19	Ancon, Panama	Male	University of Texas University of Oklahoma
97	Frederick D. Gregory	1978	8	Retired	1941-01-07	Washington, DC	Male	US Air Force Academy Georgia Washington University
98	Frederick H. Hauck	1978	8	Retired	1941-04-11	Long Beach, CA	Male	Tulane University
99	Frederick W. Sturckow	1995	15	Retired	1961-08-11	La Mesa, CA	Male	California Polytechnic State University
100	G. David Low	1984	10	Deceased	1956-02-19	Cleveland, OH	Male	Washington Lee University
101	G. Reid Wiseman	2009	20	Active	1975-11-11	Baltimore, MD	Male	Rensselaer Polytechnic Institute Johns Hopkins University
102	Garrett E. Reisman	1998	17	Retired	1968-02-10	Morristown, NJ	Male	University of Pennsylvania California Institute of Technology
103	Gary E. Payton	1998	17	Retired	1948-06-20	Rock Island, IL	Male	US Air Force Academy Purdue University

index	Name	Year	GroupNum	Status	Birth_Date	Birth_Place	Gender	Alma_Mat
104	George D. Nelson	1978	8	Retired	1950-07-13	Charles City, IA	Male	Harvey Mudd College; University of Washington
105	George D. Zamka	1998	17	Retired	1962-06-29	Jersey City, NJ	Male	US Naval Academy; Florida Institute of Technology
106	Gerald P. Carr	1966	5	Retired	1932-08-22	Denver, CO	Male	University of Southern California; US Naval Postgraduate School; Princeton University
107	Gregory B. Jarvis	1966	5	Deceased	1944-08-24	Detroit, MI	Male	State University of New York at Buffalo; Northeastern University
108	Gregory C. Johnson	1998	17	Management	1954-07-30	Seattle, WA	Male	University of Washington
109	Gregory E. Chamitoff	1998	17	Active	1962-08-06	Montreal, Canada	Male	California Polytechnic State University; California Institute of Technology; MIT
110	Gregory H. Johnson	1998	17	Active	1962-05-12	London, England	Male	US Air Force Academy; Columbia University; University of Texas
111	Gregory J. Harbaugh	1987	12	Retired	1956-04-15	Cleveland, OH	Male	Purdue University; University of Houston Clear Lake
112	Gregory T. Linteris	1987	12	Retired	1957-10-04	Demarest, NJ	Male	Princeton University; Stanford University
113	Guion S. Bluford Jr.	1978	8	Retired	1942-11-22	Philadelphia, PA	Male	Pennsylvania State University; A

index	Name	Year	GroupNum	Status	Birth_Date	Birth_Place	Gender	Alma_Mat
								Force Institu
								Technolog
								University
								Houston
								Clear Lak
114	Guy S. Gardner	1980	9	Retired	1948-01-06	Alta Vista, VA	Male	US Air Force Academ
								Purdue Universi
115	Harrison H. Schmitt	1965	4	Retired	1935-07-03	Santa Rita, NM	Male	California Institute
								Technolog
								Harvard Universi
116	Heidemarie M. Stefanyshyn-Piper	1996	16	Retired	1963-02-07	St. Paul, MN	Female	M
117	Henry W. Hartsfield Jr.	1969	7	Retired	1933-11-21	Birmingham, AL	Male	Auburn University
								University of Tennessee
118	Jack D. Fischer	2009	20	Active	1974-01-23	Louisville, CO	Male	US Air Force Academy; M
119	Jack R. Lousma	1966	5	Retired	1936-02-29	Grand Rapids, MI	Male	University of Michigan; U
								Nav Postgradua
								School
120	Jake Garn	1966	5	Retired	1932-10-12	Richfield, UT	Male	University of Utah
121	James A. Lovell Jr.	1962	2	Retired	1928-03-25	Cleveland, OH	Male	US Naval Academ
122	James A. McDivitt	1984	2	Retired	1929-06-10	Chicago, IL	Male	University of Michigan
123	James A. Pawelczyk	1984	2	Retired	1960-09-20	Buffalo, NY	Male	University of Rochester
								Pennsylvani
								State Universit
								University of North Texas
124	James B. Irwin	1966	5	Deceased	1930-03-17	Pittsburgh, PA	Male	US Naval Academ
								University of Michigan
125	James C. Adamson	1984	10	Retired	1946-03-03	Warsaw, NY	Male	US Military Academ

index	Name	Year	GroupNum	Status	Birth_Date	Birth_Place	Gender	Alma_Mate
								Princeton University
126	James D. Halsell Jr.	1990	13	Retired	1956-09-29	Monroe, LA	Male	US Air Force Academy; Troy State University; University of Air Force Institute of Technology
127	James D. van Hoften	1978	8	Retired	1944-06-11	Fresno, CA	Male	University of California Berkeley; Colorado State University
128	James D. Wetherbee	1984	10	Retired	1952-11-27	Flushing, NY	Male	University of Notre Dame
129	James F. Buchli	1978	8	Retired	1945-06-20	New Rockford, ND	Male	US Naval Academy; University of West Florida
130	James F. Reilly II	1995	15	Retired	1954-03-18	Mountain Home Air Force Base, ID	Male	University of Texas-Dallas
131	James H. Newman	1990	13	Retired	1956-10-16	San Diego, CA	Male	Dartmouth College; Rice University
132	James M. Kelly	1996	16	Management	1964-05-14	Burlington, IA	Male	US Air Force Academy; University of Alabama
133	James P. Bagian	1980	9	Retired	1952-02-22	Philadelphia, PA	Male	Drexel University; Thomas Jefferson University
134	James P. Dutton Jr.	2004	19	Management	1968-11-20	Eugene, OR	Male	US Air Force Academy; University of Washington
135	James S. Voss	1987	12	Retired	1949-03-03	Cordova, AL	Male	Auburn University; University of Colorado

index	Name	Year	GroupNum	Status	Birth_Date	Birth_Place	Gender	Alma_Mat
136	Janet L. Kavandi	1995	15	Management	1959-07-17	Springfield, MO	Female	Missouri Southern State College University of Missouri University of Washington
137	Janice E. Voss	1990	13	Deceased	1956-10-08	South Bend, IN	Female	Purdue University M
138	Jay Clark Buckley	1990	13	Retired	1956-06-06	New York, NY	Male	Cornell University
139	Jeanette J. Epps	2009	20	Active	1970-11-03	Syracuse, NY	Female	LeMoyne College University of Maryland
140	Jeffrey A. Hoffman	1978	8	Retired	1944-11-02	Brooklyn, NY	Male	Amherst College; Rice University; Harvard University
141	Jeffrey N. Williams	1996	16	Active	1958-01-18	Superior, WI	Male	US Military Academy; US Naval Postgraduate School; US Naval War College
142	Jeffrey S. Ashby	1995	15	Retired	1954-06-01	Dallas, TX	Male	University of Idaho; University of Tennessee
143	Jerome Apt III	1985	11	Retired	1949-04-18	Springfield, MA	Male	Harvard University M
144	Jerry L. Ross	1980	9	Retired	1948-01-20	Crown Point, IN	Male	Purdue University
145	Jerry M. Linenger	1992	14	Retired	1955-01-16	Mount Clemens, MI	Male	US Naval Academy; University of Southern California

index	Name	Year	GroupNum	Status	Birth_Date	Birth_Place	Gender	Alma_Mat
								University o
								Northe
								Carolin
								Wayne Stat
								Universit
								Souther
								Illino
								Universit
								Carbondal
								Floric
								Institut
								Technolog
146	Joan E. Higginbotham	1996	16	Retired	1964-08-03	Chicago, IL	Female	US Nav
147	Joe F. Edwards Jr.	1995	15	Retired	1958-02-03	Richmond, VA	Male	Academ
148	Joe H. Engle	1966	5	Retired	1932-08-26	Dickinson, KS	Male	Universit
149	John A. Llewellyn	1967	6	Retired	1933-04-22	Cardiff, Wales	Male	Universit
150	John B. Herrington	1996	16	Retired	1958-09-14	Wetumka, OK	Male	Universit
151	John D. Olivas	1998	17	Retired	1966-05-25	North Hollywood, CA	Male	Texas-El Pas
152	John E. Blaha	1980	9	Retired	1942-08-26	San Antonio, TX	Male	Universit
153	John H. Casper	1984	10	Management	1943-07-09	Greenville, SC	Male	US Air For
154	John H. Glenn Jr.	1959	1	Deceased	1921-07-18	Cambridge, OH	Male	Academ
155	John L. Phillips	1996	16	Retired	1951-04-15	Ft. Belvoir, VA	Male	Purdue
								Universit
								US Nav
								Academ
								University
								West Florid
								Universit

index	Name	Year	GroupNum	Status	Birth_Date	Birth_Place	Gender	Alma_Mat
								California-Los Angeles
156	John L. Swigert Jr.	1966	5	Deceased	1931-08-30	Denver, CO	Male	University of Colorado; Rensselaer Polytechnic Institute; University of Hartford
157	John M. Fabian	1978	8	Retired	1939-01-28	Goosecreek, TX	Male	Washington State University; University of Air Force Institute of Technology; University of Washington
158	John M. Grunsfeld	1992	14	Management	1958-10-10	Chicago, IL	Male	Michigan University of Chicago
159	John O. Creighton	1978	8	Retired	1943-04-28	Orange, TX	Male	US Naval Academy; Georgia Institute of Technology; University of Washington
160	John W. Young	1962	2	Retired	1930-09-24	San Francisco, CA	Male	Georgia Institute of Technology
161	John-David F. Bartoe	1962	2	Retired	1944-11-17	Abington, PA	Male	Lehigh University; Georgetown University
162	Jon A. McBride	1978	8	Retired	1943-08-14	Charleston, WV	Male	US Naval Postgraduate School
163	Jose M. Hernandez	2004	19	Retired	1962-08-07	French Camp, CA	Male	University of the Pacific; University of California Santa Barbara
164	Joseph M. Acaba	2004	19	Active	1967-05-17	Inglewood, CA	Male	University of California Santa Barbara; University of Arizona
165	Joseph P. Allen	1967	6	Retired	1937-06-27	Crawfordsville, IN	Male	DePaul University; Yale University

index	Name	Year	GroupNum	Status	Birth_Date	Birth_Place	Gender	Alma_Mate
166	Joseph P. Kerwin	1965	4	Retired	1932-02-19	Oak Park, IL	Male	College of the Holy Cross Northwestern University
167	Joseph R. Tanner	1992	14	Retired	1950-01-21	Danville, IL	Male	University of Illinois
168	Judith A. Resnik	1978	8	Deceased	1949-04-05	Akron, OH	Female	Carnegie Mellon University University of Maryland
169	K. Megan McArthur	2000	18	Active	1971-08-30	Honolulu, HI	Female	University of California-Los Angeles University of California-San Diego
170	Kalpana Chawla	1995	15	Deceased	1961-06-01	Karnal, India	Female	Punjabi Engineering College University of Texas at Arlington University of Colorado
171	Karen L. Nyberg	2000	18	Active	1969-10-07	Parker's Prairie, MN	Female	University of North Dakota University of Texas
172	Karl G. Henize	1967	6	Deceased	1926-10-17	Cincinnati, OH	Male	University of Virginia University of Michigan
173	Karol J. Bobko	1969	7	Retired	1937-12-23	New York, NY	Male	US Air Force Academy University of Southern California
174	Kathleen Rubins	2009	20	Active	1978-10-14	Farmington, CT	Female	University of California-San Diego Stanford University
175	Kathryn C. Thornton	1984	10	Retired	1952-08-17	Montgomery, AL	Female	Auburn University University of Virginia
176	Kathryn D. Sullivan	1978	8	Retired	1951-10-03	Patterson, NJ	Female	University of California

index	Name	Year	GroupNum	Status	Birth_Date	Birth_Place	Gender	Alma_Mat
								Santa Cru Dalhous Universi
177	Kathryn P. Hire	1995	15	Management	1959-08-26	Mobile, AL	Female	US Nav Academ Florida Stat Institut Technolog
178	Kenneth D. Bowersox	1987	12	Retired	1956-11-14	Portsmouth, VA	Male	US Nav Academ Columb Universi
179	Kenneth D. Cameron	1984	10	Retired	1949-11-29	Cleveland, OH	Male	MIT; Michigan Stat Universi
180	Kenneth D. Cockrell	1990	13	Management	1950-04-09	Austin, TX	Male	University o Texas University o West Florid
181	Kenneth S. Reightler Jr.	1987	12	Retired	1951-03-24	Patuxent River, MD	Male	US Nav Academ; U Nav Postgradua School University o Southern Californ
182	Kenneth T. Ham	1998	17	Retired	1964-12-12	Plainfield, NJ	Male	US Nav Academ; U Nav Postgradua School
183	Kent V. Rominger	1992	14	Retired	1956-08-07	Del Norte, CO	Male	Colorad Stat University; U Nav Postgradua School
184	Kevin A. Ford	2000	18	Active	1960-07-07	Portland, IN	Male	University o Notre Dam Troy Stat Universit University o Florida; A Force Instit of Technolog
185	Kevin P. Chilton	1987	12	Retired	1954-03-11	Los Angeles, CA	Male	US Air For Academ

index	Name	Year	GroupNum	Status	Birth_Date	Birth_Place	Gender	Alma_Mate
186	Kevin R. Kregel	1992	14	Retired	1956-09-16	New York, NY	Male	Columbia University
187	Kjell N. Lindgren	2009	20	Active	1973-01-23	Taipei, Taiwan	Male	US Air Force Academy, Troy State University
188	L. Blaine Hammond Jr.	1984	10	Retired	1952-01-16	Savannah, GA	Male	US Air Force Academy, Georgia Institute of Technology
189	L. Gordon Cooper Jr.	1959	1	Deceased	1927-03-06	Shawnee, OK	Male	Air Force Institute of Technology
190	Laurel B. Clark	1996	16	Deceased	1961-03-10	Ames, IA	Female	University of Wisconsin-Madison
191	Lawrence J. Delucas	1996	16	Retired	1950-07-11	Syracuse, NY	Male	University of Alabama at Birmingham
192	Lee J. Archambault	1998	17	Retired	1960-08-25	Oak Park, IL	Male	University of Illinois Urbana-Champaign
193	Lee M. Morin	1996	16	Management	1952-09-09	Manchester, NH	Male	University of New Hampshire
194	Leland D. Melvin	1998	17	Management	1964-02-15	Lynchburg, VA	Male	University of Richmond
195	Leroy Chiao	1990	13	Retired	1960-08-28	Milwaukee, WI	Male	University of California-Berkeley

index	Name	Year	GroupNum	Status	Birth_Date	Birth_Place	Gender	Alma_Mat
								California Santa Barbara
196	Linda M. Godwin	1985	11	Retired	1952-07-02	Cape Girardeau, MO	Female	Southeast Missouri State University University of Missouri
197	Lisa M. Nowak	1996	16	Retired	1963-05-10	Washington, DC	Female	US Naval Academy; US Naval Postgraduate School
198	Lodewijk van den Berg	1996	16	Retired	1932-03-24	Sluiskil, Netherlands	Male	Delaware University University of Technology University of Delaware
199	Loren J. Shriver	1978	8	Retired	1944-09-23	Jefferson, IA	Male	US Air Force Academy Purdue University
200	Loren W. Acton	1978	8	Retired	1936-03-07	Lewiston, MT	Male	Montana State University University of Colorado
201	M. Scott Carpenter	1959	1	Retired	1925-05-01	Boulder, CO	Male	University of Colorado
202	Mae C. Jemison	1987	12	Retired	1956-10-17	Decatur, AL	Female	Stanford University Cornell University
203	Margaret Rhea Seddon	1978	8	Retired	1947-11-08	Murfreesboro, TN	Female	University of California Berkeley University of Tennessee
204	Mario Runco Jr.	1987	12	Management	1952-01-26	Bronx, NY	Male	City College of New York Rutgers University
205	Mark C. Lee	1984	10	Retired	1952-08-14	Viroqua, WI	Male	US Air Force Academy; M
206	Mark E. Kelly	1996	16	Retired	1964-02-21	Orange, NJ	Male	US Merchant Marine Academy; US Naval

index	Name	Year	GroupNum	Status	Birth_Date	Birth_Place	Gender	Alma_Mate
207	Mark L. Polansky	1996	16	Retired	1956-06-02	Paterson, NJ	Male	Purdue University
208	Mark N. Brown	1984	10	Retired	1951-11-18	Valparaiso, IN	Male	Purdue University; Air Force Institute of Technology
209	Mark T. Vande Hei	2009	20	Active	1966-11-10	Falls Church, VA	Male	Saint John's University; Stanford University
210	Marsha S. Ivins	1984	10	Retired	1951-04-15	Baltimore, MD	Female	University of Colorado
211	Martin J. Fettman	1984	10	Retired	1956-12-31	Brooklyn, NY	Male	Cornell University; Colorado State University
212	Mary E. Weber	1992	14	Retired	1962-08-24	Cleveland, OH	Female	Purdue University; University of California, Berkeley
213	Mary L. Cleave	1980	9	Retired	1947-02-05	Southampton, NY	Female	Colorado State University; Utah State University
214	Michael A. Baker	1985	11	Management	1953-10-27	Memphis, TN	Male	University of Texas
215	Michael Collins	1963	3	Retired	1930-10-31	Rome, Italy	Male	US Military Academy
216	Michael E. Fossum	1998	17	Active	1957-12-19	Sioux Falls, SD	Male	Texas A&M University; Air Force Institute of Technology
217	Michael E. Lopez-Alegria	1992	14	Retired	1958-05-30	Madrid, Spain	Male	US Naval Academy; University of Houston-Clear Lake

index	Name	Year	GroupNum	Status	Birth Date	Birth Place	Gender	Alma Mater
								Nav Postgradua Scho
218	Michael J. Bloomfield	1995	15	Retired	1959-03-16	Flint, MI	Male	US Air Force Academy; Order of the Dominic Universit
219	Michael J. Foreman	1998	17	Management	1957-03-29	Columbus, OH	Male	US Naval Academy; US Navy Postgradua Scho
220	Michael J. Massimino	1996	16	Management	1962-08-19	Oceanside, NY	Male	Columbia University M
221	Michael J. McCulley	1984	10	Retired	1943-08-04	San Diego, CA	Male	Purdue Universi
222	Michael J. Smith	1980	9	Deceased	1945-04-30	Beaufort, NC	Male	US Naval Academy; US Navy Postgradua Scho
223	Michael L. Coats	1978	8	Retired	1946-01-16	Sacramento, CA	Male	US Naval Academ Georg Washington University; US Navy Postgradua Scho
224	Michael L. Gernhardt	1992	14	Management	1956-05-04	Mansfield, OH	Male	Vanderbilt University University of Pennsylvania
225	Michael P. Anderson	1995	15	Deceased	1959-12-25	Plattsburgh, NY	Male	University of Washington; Creighton Universi
226	Michael R. Barratt	2000	18	Active	1959-04-16	Vancouver, WA	Male	University of Washington; Northwestern Universit Wright State Universi
227	Michael R. Clifford	1990	13	Retired	1952-10-13	San Bernardino, CA	Male	US Military Academ Georg Institute of Technolog

index	Name	Year	GroupNum	Status	Birth_Date	Birth_Place	Gender	Alma_Mat
228	Michael S. Hopkins	2009	20	Active	1968-12-28	Lebanon, MO	Male	University of Illinois Urbana-Champaign Stanford University
229	Michael T. Good	2000	18	Management	1962-10-13	Parma, OH	Male	University of Notre Dame
230	Millie Hughes-Fulford	2000	18	Retired	1945-12-21	Mineral Wells, TX	Female	Tarleton State University University of Texas at San Antonio Woman's University
231	N. Jan Davis	1987	12	Retired	1953-11-01	Cocoa Beach, FL	Female	Georgia Institute of Technology Auburn University University of Alabama Huntsville
232	Nancy J. Currie	1990	13	Management	1958-12-29	Wilmington, DE	Female	Ohio State University University of Southern California University of Houston
233	Neil A. Armstrong	1962	2	Deceased	1930-08-05	Wapakoneta, OH	Male	Purdue University University of Southern California
234	Neil W. Woodward III	1998	17	Retired	1962-07-26	Chicago, IL	Male	Michigan University Texas-Austin Georgia Institute of Technology Washington University
235	Nicolas J. M. Patrick	1998	17	Retired	1964-03-22	North Yorkshire, England	Male	University of Cambridge Massachusetts Institute of Technology
236	Nicole P. Stott	2000	18	Active	1962-11-19	Albany, NY	Female	Embry-Riddle Aeronautical University University of Central Florida
237	Norman E. Thagard	1978	8	Retired	1943-07-03	Marianna, FL	Male	Florida State University University of Texas at Austin

index	Name	Year	GroupNum	Status	Birth_Date	Birth_Place	Gender	Alma_Mat
238	Owen K. Garriott	1965	4	Retired	1930-11-22	Enid, OK	Male	University of Oklahoma; Stanford University
239	Pamela A. Melroy	1995	15	Retired	1961-09-17	Palo Alto, CA	Female	Wellesley College; M.I.T.
240	Patrick G. Forrester	1996	16	Management	1957-03-31	El Paso, TX	Male	US Military Academy; University of Virginia
241	Paul D. Scully-Power	1996	16	Retired	1944-05-28	Sydney, Australia	Male	University of Sydney
242	Paul J. Weitz	1966	5	Retired	1932-06-25	Erie, PA	Male	Pennsylvania State University; U.S. Naval Postgraduate School
243	Paul S. Lockhart	1996	16	Retired	1956-04-28	Amarillo, TX	Male	Texas Tech University; University of Texas at Austin
244	Paul W. Richards	1996	16	Management	1964-05-20	Scranton, PA	Male	Drexel University; University of Maryland
245	Peggy A. Whitson	1996	16	Active	1960-02-09	Mt. Ayr, IA	Female	Iowa Wesleyan College; Rice University
246	Peter J. K. Wisoff	1990	13	Retired	1958-08-16	Norfolk, VA	Male	University of Virginia; Stanford University
247	Philip K. Chapman	1967	6	Retired	1935-03-05	Melbourne, Australia	Male	University of Sydney; M.I.T.
248	Pierre J. Thuot	1985	11	Retired	1955-05-19	Groton, CT	Male	US Naval Academy; University of Southern California
249	Piers J. Sellers	1996	16	Management	1955-04-11	Crowborough, England	Male	University of Edinburgh; Leeds University

index	Name	Year	GroupNum	Status	Birth_Date	Birth_Place	Gender	Alma_Mat
250	Randolph J. Bresnik	2004	19	Active	1967-09-11	Fort Knox, KY	Male	The Citadel University of Tennessee Tennessee Knoxville
251	Rex J. Walheim	1996	16	Active	1962-10-10	Redwood, CA	Male	University of California Berkeley University of Houston
252	Richard A. Mastracchio	1996	16	Active	1960-02-11	Waterbury, CT	Male	University of Connecticut Rensselaer Polytechnic Institute University of Houston Clear Lake
253	Richard A. Searfoss	1990	13	Retired	1956-06-05	Mount Clemens, MI	Male	US Air Force Academy California Institute of Technology
254	Richard F. Gordon Jr.	1963	3	Retired	1929-10-05	Seattle, WA	Male	University of Washington
255	Richard H. Truly	1969	7	Retired	1937-11-12	Fayette, MS	Male	Georgia Institute of Technology
256	Richard J. Hieb	1985	11	Retired	1955-09-21	Jamestown, ND	Male	Northwestern Nazarene College University of Colorado
257	Richard M. Linnehan	1992	14	Management	1957-09-19	Lowell, MA	Male	University of New Hampshire Ohio State University Harvard University
258	Richard M. Mullane	1978	8	Retired	1945-09-10	Wichita Falls, TX	Male	US Military Academy; US Air Force Institute of Technology
259	Richard N. Richards	1980	9	Retired	1946-08-24	Key West, FL	Male	University of Missouri University of West Florida
260	Richard O. Covey	1978	8	Retired	1946-08-01	Fayetteville, AR	Male	US Air Force Academy

index	Name	Year	GroupNum	Status	Birth_Date	Birth_Place	Gender	Alma_Mat
261	Richard R. Arnold II	2004	19	Active	1963-11-26	Cheverly, MD	Male	Purdue University
262	Rick D. Husband	1995	15	Deceased	1957-07-12	Amarillo, TX	Male	Frostburg State University
263	Robert A. Parker	1967	6	Retired	1936-12-14	New York, NY	Male	Amherst College
264	Robert C. Springer	1980	9	Retired	1942-05-21	St. Louis, MO	Male	US Naval Academy; US Naval Postgraduate School
265	Robert Cenker	1980	9	Retired	1948-11-05	Uniontown, PA	Male	Pennsylvania State University
266	Robert D. Cabana	1985	11	Management	1949-01-23	Minneapolis, MN	Male	Rutgers University
267	Robert F. Overmyer	1969	7	Deceased	1936-07-14	Lorain, OH	Male	Baldwin Wallace College; US Naval Postgraduate School
268	Robert L. Behnken	2000	18	Active	1970-07-28	Creve Coeur, MO	Male	Washington University
269	Robert L. Crippen	1969	7	Retired	1937-09-11	Beaumont, TX	Male	University of Texas
270	Robert L. Curbeam Jr.	1995	15	Retired	1962-03-05	Baltimore, MD	Male	US Naval Academy; US Naval

index	Name	Year	GroupNum	Status	Birth_Date	Birth_Place	Gender	Alma_Mat
271	Robert L. Gibson	1978	8	Retired	1946-10-30	Cooperstown, NY	Male	California Polytechnic Institute
272	Robert L. Satcher Jr.	2004	19	Retired	1965-09-22	Hampton, VA	Male	MIT; Harvard University
273	Robert L. Stewart	1978	8	Retired	1942-08-13	Washington, DC	Male	University of Southern Mississippi; University of Texas at Arlington
274	Robert Shane Kimbrough	2004	19	Active	1967-06-04	Killeen, TX	Male	US Military Academy; Georgia Institute of Technology
275	Roger B. Chaffee	1963	3	Deceased	1935-02-15	Grand Rapids, MI	Male	Purdue University
276	Roger K. Crouch	1963	3	Retired	1940-09-12	Jamestown, TN	Male	Tennessee Polytechnic Institute; Virgin Polytechnic Institute
277	Ronald A. Parise	1963	3	Deceased	1951-05-24	Warren, OH	Male	Youngstown State University; University of Florida
278	Ronald E. Evans Jr.	1966	5	Deceased	1933-11-10	St. Francis, KS	Male	University of Kansas; US Naval Postgraduate School
279	Ronald E. McNair	1978	8	Deceased	1950-10-21	Lake City, SC	Male	North Carolina A&T State College
280	Ronald J. Garan Jr.	2000	18	Management	1961-10-20	Yonkers, NY	Male	Stanford University; New York Embry-Riddle Aeronautical University

index	Name	Year	GroupNum	Status	Birth_Date	Birth_Place	Gender	Alma_Mat
281	Ronald J. Grabe	1980	9	Retired	1945-06-13	New York, NY	Male	University of Florida
282	Ronald M. Sega	1990	13	Retired	1952-12-04	Cleveland, OH	Male	US Air Force Academy
283	Roy D. Bridges Jr.	1980	9	Retired	1943-07-19	Atlanta, GA	Male	US Air Force Academy
284	Russell L. Schweickart	1963	3	Retired	1935-10-25	Neptune, NJ	Male	M
285	S. Christa McAuliffe	1963	3	Deceased	1948-09-02	Boston, MA	Female	Framingham State College
286	S. David Griggs	1978	8	Deceased	1939-09-07	Portland, OR	Male	US Naval Academy
287	Sally K. Ride	1978	8	Deceased	1951-05-26	Los Angeles, CA	Female	Stanford University
288	Samuel T. Durrance	1978	8	Retired	1943-09-17	Tallahassee, FL	Male	California State University
289	Sandra H. Magnus	1996	16	Retired	1964-10-30	Belleville, IL	Female	University of Missouri-Rolla; Georgia Institute of Technology
290	Scott D. Altman	1995	15	Retired	1959-08-15	Lincoln, IL	Male	University of Illinois; US Naval Postgraduate School
291	Scott D. Tingle	2009	20	Active	1965-07-19	Attleboro, MA	Male	Southeastern Massachusetts University

index	Name	Year	GroupNum	Status	Birth_Date	Birth_Place	Gender	Alma_Mat
292	Scott E. Parazynski	1992	14	Retired	1961-07-28	Little Rock, AR	Male	Stanford University
293	Scott J. Horowitz	1992	14	Retired	1957-03-24	Philadelphia, PA	Male	California State University Northridge Georgia Institute of Technology
294	Scott J. Kelly	1996	16	Active	1964-02-21	Orange, NJ	Male	State University of New York Maritime College University of Tennessee Knoxville
295	Serena M. Aunon	2009	20	Active	1976-04-09	Indianapolis, IN	Female	Georgia Washington University University of Texas
296	Shannon W. Lucid	1978	8	Retired	1943-01-14	Shanghai, China	Female	University of Oklahoma
297	Shannon Walker	2004	19	Active	1965-06-04	Houston, TX	Female	Rice University
298	Sherwood C. Spring	1980	9	Retired	1944-09-23	Hartford, CT	Male	US Military Academy University of Arizona
299	Sidney M. Gutierrez	1984	10	Retired	1951-06-27	Albuquerque, NM	Male	US Air Force Academy Webster College
300	Stanley G. Love	1998	17	Management	1965-06-08	San Diego, CA	Male	Harvey Mudd College University of Washington
301	Stephanie D. Wilson	1996	16	Active	1966-09-27	Boston, MA	Female	Harvard University University of Texas
302	Stephen G. Bowen	2000	18	Active	1964-02-13	Cohasset, MA	Male	US Naval Academy; M

index	Name	Year	GroupNum	Status	Birth_Date	Birth_Place	Gender	Alma_Mate
303	Stephen K. Robinson	1995	15	Retired	1955-10-26	Sacramento, CA	Male	University of California, Davis; Stanford University
304	Stephen N. Frick	1996	16	Management	1961-09-30	Pittsburgh, PA	Male	US Naval Academy; US Naval Postgraduate School
305	Stephen S. Oswald	1985	11	Retired	1951-06-30	Seattle, WA	Male	US Naval Academy
306	Steven A. Hawley	1978	8	Retired	1951-12-12	Ottawa, KS	Male	University of Kansas; University of California
307	Steven L. Smith	1992	14	Management	1958-12-30	Phoenix, AZ	Male	Stanford University
308	Steven R. Nagel	1978	8	Retired	1946-10-27	Canton, IL	Male	University of Illinois at Urbana-Champaign; California State University, Fresno
309	Steven R. Swanson	1998	17	Active	1960-12-03	Syracuse, NY	Male	University of Colorado Boulder; Florida Institute of Technology; Atlanta University Center; Texas A&M University
310	Steven W. Lindsey	1995	15	Retired	1960-08-24	Arcadia, CA	Male	US Air Force Academy; US Air Force Institute of Technology
311	Story Musgrave	1967	6	Retired	1935-08-19	Boston, MA	Male	Syracuse University; Marietta College; University of California-Los Angeles; University of Kentucky; University of Houston

index	Name	Year	GroupNum	Status	Birth_Date	Birth_Place	Gender	Alma_Mat
								Columb Universi
312	Stuart A. Roosa	1966	5	Deceased	1933-08- 16	Durango, CO	Male	University of Colorado
313	Sunita L. Williams	1998	17	Active	1965-09- 19	Euclid, OH	Female	US Nav Academ Floric Institut Technolog
314	Susan J. Helms	1990	13	Retired	1958-02- 26	Charlotte, NC	Female	US Air Force Academ Stanfor Universi
315	Susan L. Kilrain (Still)	1995	15	Retired	1961-10- 24	Augusta, GA	Female	Embry-Ridd Universit Georg Institut Technolog
316	Tamara E. Jernigan	1985	11	Retired	1959-05- 07	Chattanooga, TN	Female	Stanfor Universit University of Califor Berkeley; Ric Universi
317	Taylor G. Wang	1985	11	Retired	1940-06- 16	Jiangxi, China	Male	University of California Los Angeles
318	Terence T. Henricks	1985	11	Retired	1952-07- 05	Bryan, OH	Male	US Air Force Academ Golden Ga Universi
319	Terrence W. Wilcutt	1990	13	Management	1949-10- 31	Russellville, KY	Male	Wester Kentuck Universi
320	Terry J. Hart	1978	8	Retired	1946-10- 27	Pittsburgh, PA	Male	Lehic Universit MIT; Rutge Universi
321	Terry W. Virts Jr.	2000	18	Active	1967-12- 01	Baltimore, MD	Male	US Air Force Academ Embry-Ridd Aeronautic Universi
322	Theodore C. Freeman	1963	3	Deceased	1930-02- 18	Haversford, PA	Male	US Nav Academ

index	Name	Year	GroupNum	Status	Birth_Date	Birth_Place	Gender	Alma_Mat
323	Thomas D. Akers	1987	12	Retired	1951-05-20	St. Louis, MO	Male	University of Michigan
324	Thomas D. Jones	1990	13	Retired	1955-01-22	Baltimore, MD	Male	US Air Force Academy; University of Arizona
325	Thomas H. Marshburn	2004	19	Active	1960-08-29	Statesville, NC	Male	Davidson College; University of Virginia; Wake Forest University; University of Texas Medical Branch Galveston
326	Thomas K. Mattingly II	1966	5	Retired	1936-03-17	Chicago, IL	Male	Auburn University
327	Thomas P. Stafford	1962	2	Retired	1930-09-17	Weatherford, OK	Male	US Naval Academy
328	Timothy J. Creamer	1998	17	Management	1959-11-15	Ft. Huachuca, AZ	Male	Loyola Marymount College; M
329	Timothy L. Kopra	2000	18	Active	1963-04-09	Austin, TX	Male	US Military Academy; Georgia Institute of Technology; US Army War College
330	Tracy E. Caldwell (Dyson)	1998	17	Active	1969-08-14	Arcadia, CA	Female	California State University Fullerton; University of California Davis
331	Vance D. Brand	1966	5	Retired	1931-05-09	Longmont, CA	Male	University of Colorado Boulder; University of California Los Angeles
332	Virgil I. Grissom	1959	1	Deceased	1926-04-03	Mitchell, IN	Male	Purdue University

index	Name	Year	GroupNum	Status	Birth_Date	Birth_Place	Gender	Alma_Mat
333	Walter Cunningham	1963	3	Retired	1932-03-16	Creston, IA	Male	University of California-Los Angeles
334	Walter M. Schirra Jr.	1959	1	Deceased	1923-03-12	Hackensack, NJ	Male	US Naval Academy
335	Wendy B. Lawrence	1992	14	Retired	1959-07-02	Jacksonville, FL	Female	US Naval Academy; M
336	William A. Anders	1963	3	Retired	1933-10-17	Hong Kong	Male	US Naval Academy; Air Force Institute of Technology
337	William A. Oefelein	1998	17	Retired	1965-03-29	Ft. Belvoir, VA	Male	Oregon State University; University of Tennessee
338	William A. Pyles	1998	17	Retired	1952-06-26	Hackensack, NJ	Male	US Air Force Academy; Texas A&M University
339	William B. Lenoir	1967	6	Deceased	1939-03-14	Miami, FL	Male	M
340	William C. McCool	1996	16	Deceased	1961-09-23	San Diego, CA	Male	US Naval Academy; University of Maryland; US Naval Postgraduate School
341	William E. Thornton	1967	6	Retired	1929-04-14	Faison, NC	Male	University of North Carolina
342	William F. Fisher	1980	9	Retired	1946-04-01	Dallas, TX	Male	Stanford University; University of Houston; University of Florida
343	William F. Readdy	1987	12	Retired	1952-01-24	Quonset Point, RI	Male	US Naval Academy
344	William G. Gregory	1990	13	Retired	1957-05-14	Lockport, NY	Male	US Air Force Academy; Columbia University; Troy State University
345	William M. Shepherd	1984	10	Retired	1949-07-26	Oak Ridge, TN	Male	US Naval Academy; M

index	Name	Year	GroupNum	Status	Birth_Date	Birth_Place	Gender	Alma_Mat
346	William R. Pogue	1966	5	Retired	1930-01-23	Okemah, OK	Male	Oklahoma Bapti Universit Oklahoma Stat Universi
347	William S. McArthur Jr.	1990	3	Management	1951-07-26	Laurinburg, NC	Male	US Milita Academ Georg Institute Technolog
348	Winston E. Scott	1992	14	Retired	1950-08-06	Miami, FL	Male	Florida Stat University; U Nav Postgraduati Scho
349	Yvonne D. Cagle	1996	16	Management	1959-04-24	West Point, NY	Female	San Francisc Stat Universi

This code block retrieves all of the data from the astronauts table. This could be useful for getting a general overview of all of the astronauts in the database, or for identifying specific astronauts who meet certain criteria.

In [34]: `%sql SELECT Status, COUNT(*) AS Number FROM astronauts GROUP BY Status;`

```
* sqlite:///my_database.sqlite
Done.
```

Out[34]:

Status	Number
Active	50
Deceased	48
Management	36
Retired	216

This code block groups the astronauts by their status and counts the number of astronauts in each group. This could be useful for understanding how many astronauts are active, retired, or deceased.

In [35]: `%sql SELECT Military_Branch, COUNT(*) AS Number FROM astronauts GROUP BY Military_Bran`

```
* sqlite:///my_database.sqlite
Done.
```

Out[35]:

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

This code block groups the astronauts by their military branch and counts the number of astronauts in each group. This could be useful for understanding which military branches are most represented among astronauts.

In [36]:

```
%sql SELECT Military_Rank, COUNT(*) AS Number FROM astronauts GROUP BY Military_Rank
* sqlite:///my_database.sqlite
Done.
```

Out[36]:

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

This code block groups the astronauts by their military rank and counts the number of astronauts in each group, ordered by the number of astronauts in each group in descending order. This could be useful for identifying the most common military ranks among astronauts.

In [37]:

```
%sql SELECT Gender, COUNT(*) AS Number FROM astronauts GROUP BY Gender;
* sqlite:///my_database.sqlite
Done.
```

Out[37]: Gender Number

Female	49
Male	301

This code block groups the astronauts by their gender and counts the number of astronauts in each group. This could be useful for understanding the gender distribution among astronauts.

In [97]:

```
%%sql
SELECT ROUND(AVG(life_Expectancy)) AS Average_Life_Expectancy
FROM (
    SELECT
        CASE
            WHEN Status = 'Deceased' THEN strftime('%Y', Death_Date) - strftime('%Y',
                ELSE 2023 - strftime('%Y', Birth_Date)
            END AS life_Expectancy
        FROM astronauts
    ) AS LifeExpectancySubquery;

* sqlite:///my_database.sqlite
Done.
```

Out[97]: Average_Life_Expectancy

69.0

This code block calculates the average life expectancy of all astronauts. This could be useful for understanding the overall health and well-being of astronauts.

In [99]:

```
%%sql
SELECT ROUND(AVG(Female_life_Expectancy)) AS Female_Average_Life_Expectancy
FROM (
    SELECT
        CASE
            WHEN Status = 'Deceased' AND Gender = 'Female' THEN strftime('%Y', Death_D
            WHEN Status != 'Deceased' AND Gender = 'Female' THEN 2023 - strftime('%Y',
                ELSE NULL
            END AS Female_life_Expectancy
        FROM astronauts
    ) AS FemaleLifeExpectancySubquery;

* sqlite:///my_database.sqlite
Done.
```

Out[99]: Female_Average_Life_Expectancy

62.0

This code block calculates the average life expectancy of female astronauts. This could be useful for understanding the gender differences in life expectancy among astronauts.

In [101...]

```
%%sql
SELECT ROUND(AVG(Male_life_Expectancy)) AS Male_Average_Life_Expectancy
FROM (
    SELECT
        CASE
            WHEN Status = 'Deceased' AND Gender = 'Male' THEN strftime('%Y', Death_Dat
```

```

        WHEN Status != 'Deceased' AND Gender = 'Male' THEN 2023 - strftime('%Y', BirthDate)
        ELSE NULL
    END AS Male_life_Expectancy
FROM astronauts
) AS MaleLifeExpectancySubquery;

* sqlite:///my_database.sqlite
Done.

```

Out[101]: **Male_Average_Life_Expectancy**

70.0

This code block calculates the average life expectancy of male astronauts. This could be useful for understanding the gender differences in life expectancy among astronauts.

In [54]:

```

%%sql
SELECT Graduate_Major, COUNT(*) AS Number
FROM astronauts
GROUP BY Graduate_Major
ORDER BY Number DESC
LIMIT 10;

```

```

* sqlite:///my_database.sqlite
Done.

```

Out[54]:

Graduate_Major Number

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

This code block groups the astronauts by their graduate major and counts the number of astronauts in each group, ordered by the number of astronauts in each group in descending order. This could be useful for identifying the most common graduate majors among astronauts.

In [55]:

```

%%sql
SELECT
    COUNT(*) AS Number_of_Astronauts,
    SUM(CASE WHEN Undergraduate_Major IS NOT NULL THEN 1 ELSE 0 END) AS Astronauts_with_Undergrad,
    SUM(CASE WHEN Graduate_Major IS NOT NULL THEN 1 ELSE 0 END) AS Astronauts_with_Graduate
FROM astronauts;

* sqlite:///my_database.sqlite
Done.

```

Out[55]: Number_of_Astronauts Astronauts_with_Undergraduate_Degrees Astronauts_with_Graduate_Degrees

Number_of_Astronauts	Astronauts_with_Undergraduate_Degrees	Astronauts_with_Graduate_Degrees
350	350	350

This code block counts the total number of astronauts, the number of astronauts with undergraduate degrees, and the number of astronauts with graduate degrees. This could be useful for understanding the educational background of astronauts.

In [60]:

```
%%sql
SELECT
CASE
    WHEN INSTR(Birth_Place, ',') > 0 THEN SUBSTR(Birth_Place, INSTR(Birth_Place, ','), )
    ELSE Birth_Place
END AS STATE,
COUNT(*) AS Number
FROM astronauts
GROUP BY STATE
ORDER BY Number DESC
LIMIT 5;
```

* sqlite:///my_database.sqlite

Done.

Out[60]: STATE Number

NY	30
CA	25
TX	23
OH	21
PA	19

This code block groups the astronauts by the state in which they were born and counts the number of astronauts in each group, ordered by the number of astronauts in each group in descending order. This could be useful for identifying the states that have produced the most astronauts.

In [61]:

```
%%sql
SELECT
ROUND(AVG(Space_Flights), 2) AS Average_Number_Of_Space_Flight,
ROUND(AVG(Space_Walks), 2) AS Average_Number_Of_Space_Walks
FROM astronauts;
```

* sqlite:///my_database.sqlite

Done.

Out[61]: Average_Number_Of_Space_Flight Average_Number_Of_Space_Walks

Average_Number_Of_Space_Flight	Average_Number_Of_Space_Walks
2.4	1.27

This code block calculates the average number of space flights and space walks for all astronauts. This could be useful for understanding the experience level of astronauts.

We Have generated a variety of insights about the astronauts in the database, that is:

1. Understand the overall demographics of the astronaut population, such as the gender distribution, military branch representation, and educational background.
2. Identify the most common military ranks and graduate majors among astronauts.
3. Understand the average life expectancy of astronauts, and any gender differences in life expectancy.
4. Identify the states that have produced the most astronauts.
5. Understand the average number of space flights and space walks for astronauts.

This information could be used to make informed decisions about the astronaut selection process, training programs, and other aspects of the space program.

```
In [92]: import matplotlib.pyplot as plt
```

Now, as we did data analysis in SQL we will now make visualizations based on our data:

1. Above we imported matplotlib.pyplot, Matplotlib is a comprehensive library for creating static, animated, and interactive visualizations in Python. Matplotlib makes easy things easy and hard things possible. Pyplot provides matplotlib with two key features:
 - A. MATLAB-style interface, which allows those familiar with MATLAB to adopt Python more easily.
 - B. Statefulness, which means that pyplot stores the state of an object when you first plot it. This is essential for use in the same loop or session state until plt.
2. Upload CSV file 'astronauts' to dataframe 'df' so it can be used in an easy way.
3. Import Seaborn library, Seaborn is a library for making statistical graphics in Python. It builds on top of matplotlib and integrates closely with pandas data structures. Seaborn helps you explore and understand your data.

```
In [68]: df = pd.read_csv(r'C:\Users\hp\Downloads\astronauts.csv')
```

```
In [69]: import seaborn as sns
```

```
In [70]: df.head()
```

Out[70]:

	Name	Year	GroupNum	Status	Birth_Date	Birth_Place	Gender	Alma_Mater	Undergraduate_Major	Military_Rank	Military_Branch	Space_Flights	Space_Flight_hr	Space_Walks	Space_Walks_hr	Missions	Death_Date	Death_Mission
0	Alan B. Shepard Jr.	1959	1	Deceased	1923-11-18	East Derry, NH	Male	US Naval Academy	Navy	Captain	US Navy	5	100	2	0.0	15	1998-07-19	2009-01-06
1	Alan G. Poindexter	1998	17	Deceased	1961-11-05	Pasadena, CA	Male	Georgia Institute of Technology; US Naval Post...	Aerospace Engin...	Commander	US Navy	1	100	1	0.0	1998-07-19	2009-01-06	
2	Alan L. Bean	1963	3	Deceased	1932-03-15	Wheeler, TX	Male	University of Texas	Aerospace Engin...	Colonel	US Air Force	1	100	1	0.0	1963-07-19	2009-01-06	
3	Albert Sacco Jr.	1963	3	Retired	1949-05-03	Boston, MA	Male	Northeastern University; MIT	Chemical Engin...	Colonel	US Army	1	100	1	0.0	1963-07-19	2009-01-06	
4	Alfred M. Worden	1966	5	Retired	1932-02-07	Jackson, MI	Male	US Military Academy; University of Michigan	Military	Colonel	US Air Force	1	100	1	0.0	1966-07-19	2009-01-06	

In [71]:

df.info()

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 350 entries, 0 to 349
Data columns (total 19 columns):
 #   Column           Non-Null Count  Dtype  
--- 
 0   Name             350 non-null    object  
 1   Year              350 non-null    int64  
 2   GroupNum          350 non-null    int64  
 3   Status             350 non-null    object  
 4   Birth_Date        350 non-null    object  
 5   Birth_Place       350 non-null    object  
 6   Gender             350 non-null    object  
 7   Alma_Mater         350 non-null    object  
 8   Undergraduate_Major 350 non-null    object  
 9   Graduate_Major     350 non-null    object  
 10  Military_Rank      350 non-null    object  
 11  Military_Branch    350 non-null    object  
 12  Space_Flights       350 non-null    int64  
 13  Space_Flight_hr     350 non-null    int64  
 14  Space_Walks         350 non-null    int64  
 15  Space_Walks_hr      350 non-null    float64 
 16  Missions            350 non-null    object  
 17  Death_Date          48 non-null     object  
 18  Death_Mission       48 non-null     object  
dtypes: float64(1), int64(5), object(13)
memory usage: 52.1+ KB
```

In [72]:

df.describe()

Out[72]:

	Year	GroupNum	Space_Flights	Space_Flight_hr	Space_Walks	Space_Walks_hr
count	350.000000	350.000000	350.000000	350.000000	350.000000	350.000000
mean	1984.682857	11.225714	2.397143	1272.054286	1.271429	7.861429
std	13.369104	5.229180	1.416033	1908.599287	2.069851	13.456289
min	1959.000000	1.000000	0.000000	0.000000	0.000000	0.000000
25%	1978.000000	8.000000	1.000000	296.500000	0.000000	0.000000
50%	1987.000000	12.000000	2.000000	608.500000	0.000000	0.000000
75%	1996.000000	16.000000	3.000000	1118.750000	2.000000	12.000000
max	2009.000000	20.000000	7.000000	12818.000000	10.000000	67.000000

In [73]:

df.isna()

Out[73]:

	Name	Year	GroupNum	Status	Birth_Date	Birth_Place	Gender	Alma_Mater	Undergraduate_M
0	False	False		False	False	False	False	False	False
1	False	False		False	False	False	False	False	False
2	False	False		False	False	False	False	False	False
3	False	False		False	False	False	False	False	False
4	False	False		False	False	False	False	False	False
...
345	False	False		False	False	False	False	False	False
346	False	False		False	False	False	False	False	False
347	False	False		False	False	False	False	False	False
348	False	False		False	False	False	False	False	False
349	False	False		False	False	False	False	False	False

350 rows × 19 columns

In [74]:

df.isna().sum()

```
Out[74]: Name      0  
Year      0  
GroupNum  0  
Status     0  
Birth_Date 0  
Birth_Place 0  
Gender     0  
Alma_Mater 0  
Undergraduate_Major 0  
Graduate_Major 0  
Military_Rank 0  
Military_Branch 0  
Space_Flights 0  
Space_Flight_hr 0  
Space_Walks 0  
Space_Walks_hr 0  
Missions    0  
Death_Date   302  
Death_Mission 302  
dtype: int64
```

```
In [75]: df.dropna()
```

Out[75]:

	Name	Year	GroupNum	Status	Birth_Date	Birth_Place	Gender	Alma_Mater	Underg
0	Alan B. Shepard Jr.	1959	1	Deceased	1923-11-18	East Derry, NH	Male	US Naval Academy	
1	Alan G. Poindexter	1998	17	Deceased	1961-11-05	Pasadena, CA	Male	Georgia Institute of Technology; US Naval Post...	Aerospace
2	Alan L. Bean	1963	3	Deceased	1932-03-15	Wheeler, TX	Male	University of Texas	
19	Brian T. O'Leary	1967	6	Deceased	1940-01-27	Boston, MA	Male	Williams College; Georgetown University; Unive...	
31	Charles A. Bassett II	1963	3	Deceased	1931-12-30	Dayton, OH	Male	Texas Technological College	Electri
32	Charles Conrad Jr.	1962	2	Deceased	1930-05-02	Philadelphia, PA	Male	Princeton University	
35	Charles E. Brady Jr.	1992	14	Deceased	1951-08-12	Pinehurst, NC	Male	University of North Carolina at Chapel Hill; D...	
39	Charles Lacy Veach	1984	10	Deceased	1944-09-18	Chicago, IL	Male	US Air Force Academy	
46	Clifton C. Williams Jr.	1963	3	Deceased	1932-09-26	Mobile, AL	Male	Auburn University	
57	David M. Brown	1996	16	Deceased	1956-04-16	Arlington, VA	Male	College of William & Mary; Eastern Virginia Me...	
58	David M. Walker	1978	8	Deceased	1944-05-20	Columbus, GA	Male	US Naval Academy	
66	Donald K. Slayton	1959	1	Deceased	1924-03-01	Sparta, WI	Male	University of Minnesota	
70	Donn F. Eisele	1963	3	Deceased	1930-05-23	Columbus, OH	Male	US Naval Academy; US Air Force Institute of Te...	
79	Edward G. Givens Jr.	1966	5	Deceased	1930-01-05	Quanah, TX	Male	US Naval Academy	

	Name	Year	GroupNum	Status	Birth_Date	Birth_Place	Gender	Alma_Mater	Undergr
80	Edward H. White II	1962	2	Deceased	1930-11-14	San Antonio, TX	Male	US Military Academy; University of Michigan	
85	Elliot M. See Jr.	1962	2	Deceased	1927-07-23	Dallas, TX	Male	US Merchant Marine Academy; University of Cali...	
86	Ellison S. Onizuka	1978	8	Deceased	1946-06-24	Kealakekua, HI	Male	University of Colorado	Aerospace
88	Eugene A. Cernan	1963	3	Deceased	1934-03-14	Chicago, IL	Male	Purdue University; US Naval Postgraduate School	Electrical
91	Francis R. Scobee	1978	8	Deceased	1939-05-19	Cle Elum, WA	Male	University of Arizona	Aerospace
100	G. David Low	1984	10	Deceased	1956-02-19	Cleveland, OH	Male	Washington & Lee University; Cornell University...	Physics
107	Gregory B. Jarvis	1966	5	Deceased	1944-08-24	Detroit, MI	Male	State University of New York at Buffalo; North...	Electrical
124	James B. Irwin	1966	5	Deceased	1930-03-17	Pittsburgh, PA	Male	US Naval Academy; University of Michigan	
137	Janice E. Voss	1990	13	Deceased	1956-10-08	South Bend, IN	Female	Purdue University; MIT	Engineering
154	John H. Glenn Jr.	1959	1	Deceased	1921-07-18	Cambridge, OH	Male	Muskingum College	
156	John L. Swigert Jr.	1966	5	Deceased	1931-08-30	Denver, CO	Male	University of Colorado; Rensselaer Polytechnic...	
168	Judith A. Resnik	1978	8	Deceased	1949-04-05	Akron, OH	Female	Carnegie-Mellon University; University of Mary...	Electrical
170	Kalpana Chawla	1995	15	Deceased	1961-06-01	Karnal, India	Female	Punjab Engineering	

	Name	Year	GroupNum	Status	Birth_Date	Birth_Place	Gender	Alma_Mater	Underg
172	Karl G. Henize	1967	6	Deceased	1926-10-17	Cincinnati, OH	Male	University of Virginia; University of Michigan	College; University of Texa...
189	L. Gordon Cooper Jr.	1959	1	Deceased	1927-03-06	Shawnee, OK	Male	Air Force Institute of Technology	
190	Laurel B. Clark	1996	16	Deceased	1961-03-10	Ames, IA	Female	University of Wisconsin-Madison	
222	Michael J. Smith	1980	9	Deceased	1945-04-30	Beaufort, NC	Male	US Naval Academy; US Naval Postgraduate School	
225	Michael P. Anderson	1995	15	Deceased	1959-12-25	Plattsburgh, NY	Male	University of Washington; Creighton University	Physic
233	Neil A. Armstrong	1962	2	Deceased	1930-08-05	Wapakoneta, OH	Male	Purdue University; University of Southern Calif...	
262	Rick D. Husband	1995	15	Deceased	1957-07-12	Amarillo, TX	Male	Texas Tech University; California State Univer...	
267	Robert F. Overmyer	1969	7	Deceased	1936-07-14	Lorain, OH	Male	Baldwin Wallace College; US Naval Postgraduate...	
275	Roger B. Chaffee	1963	3	Deceased	1935-02-15	Grand Rapids, MI	Male	Purdue University	
277	Ronald A. Parise	1963	3	Deceased	1951-05-24	Warren, OH	Male	Youngstown State University; University of Flo...	
278	Ronald E. Evans Jr.	1966	5	Deceased	1933-11-10	St. Francis, KS	Male	University of Kansas; US Naval Postgraduate Sc...	Electri
279	Ronald E. McNair	1978	8	Deceased	1950-10-21	Lake City, SC	Male	North Carolina A&T State	

	Name	Year	GroupNum	Status	Birth_Date	Birth_Place	Gender	Alma_Mater	Underg
285	S. Christa McAuliffe	1963	3	Deceased	1948-09-02	Boston, MA	Female	College; MIT Framingham State College; Bowie State College	
286	S. David Griggs	1978	8	Deceased	1939-09-07	Portland, OR	Male	US Naval Academy; George Washington University	
287	Sally K. Ride	1978	8	Deceased	1951-05-26	Los Angeles, CA	Female	Stanford University	
312	Stuart A. Roosa	1966	5	Deceased	1933-08-16	Durango, CO	Male	University of Colorado	
322	Theodore C. Freeman	1963	3	Deceased	1930-02-18	Haversford, PA	Male	US Naval Academy; University of Michigan	
332	Virgil I. Grissom	1959	1	Deceased	1926-04-03	Mitchell, IN	Male	Purdue University	
334	Walter M. Schirra Jr.	1959	1	Deceased	1923-03-12	Hackensack, NJ	Male	US Naval Academy	
339	William B. Lenoir	1967	6	Deceased	1939-03-14	Miami, FL	Male	MIT	Electri
340	William C. McCool	1996	16	Deceased	1961-09-23	San Diego, CA	Male	US Naval Academy; University of Maryland; US N...	

In [76]: df.isna().sum()

```
Out[76]: Name      0  
Year      0  
GroupNum  0  
Status     0  
Birth_Date 0  
Birth_Place 0  
Gender     0  
Alma_Mater 0  
Undergraduate_Major 0  
Graduate_Major 0  
Military_Rank 0  
Military_Branch 0  
Space_Flights 0  
Space_Flight_hr 0  
Space_Walks 0  
Space_Walks_hr 0  
Missions    0  
Death_Date   302  
Death_Mission 302  
dtype: int64
```

```
In [77]: df
```

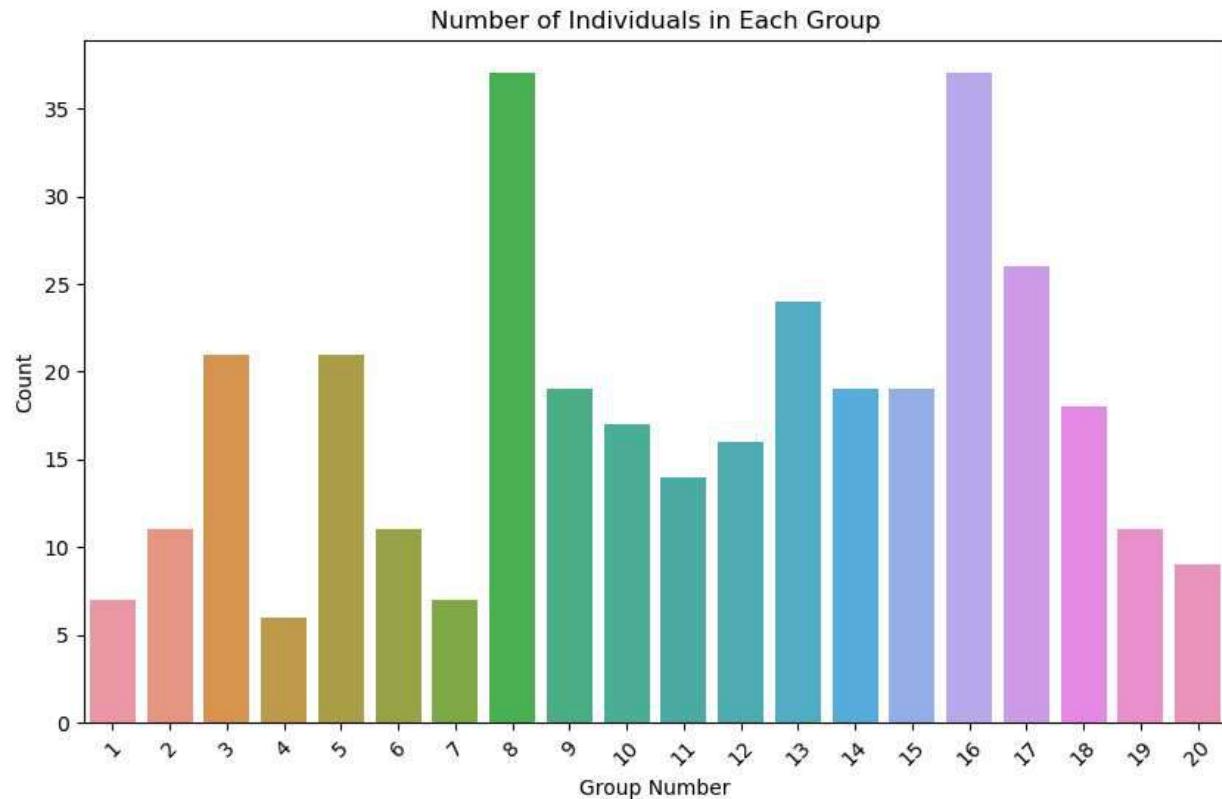
Out[77]:

	Name	Year	GroupNum	Status	Birth_Date	Birth_Place	Gender	Alma_Mater	Under
0	Alan B. Shepard Jr.	1959	1	Deceased	1923-11-18	East Derry, NH	Male	US Naval Academy	
1	Alan G. Poindexter	1998	17	Deceased	1961-11-05	Pasadena, CA	Male	Georgia Institute of Technology; US Naval Post...	Aeros...
2	Alan L. Bean	1963	3	Deceased	1932-03-15	Wheeler, TX	Male	University of Texas	
3	Albert Sacco Jr.	1963	3	Retired	1949-05-03	Boston, MA	Male	Northeastern University; MIT	Cher...
4	Alfred M. Worden	1966	5	Retired	1932-02-07	Jackson, MI	Male	US Military Academy; University of Michigan	
...
345	William M. Shepherd	1984	10	Retired	1949-07-26	Oak Ridge, TN	Male	US Naval Academy; MIT	Aeros...
346	William R. Pogue	1966	5	Retired	1930-01-23	Okemah, OK	Male	Oklahoma Baptist University; Oklahoma State Un...	
347	William S. McArthur Jr.	1990	3	Management	1951-07-26	Laurinburg, NC	Male	US Military Academy; Georgia Institute of Tech...	A
348	Winston E. Scott	1992	14	Retired	1950-08-06	Miami, FL	Male	Florida State University; US Naval Postgraduat...	
349	Yvonne D. Cagle	1996	16	Management	1959-04-24	West Point, NY	Female	San Francisco State University	

350 rows × 19 columns

In [78]: `group_counts = df['GroupNum'].value_counts()`In [79]: `plt.figure(figsize=(10, 6))
sns.barplot(x=group_counts.index, y=group_counts.values)`

```
plt.title('Number of Individuals in Each Group')
plt.xlabel('Group Number')
plt.ylabel('Count')
plt.xticks(rotation=45)
plt.show()
```



Categorical Data Analysis and Visualization:

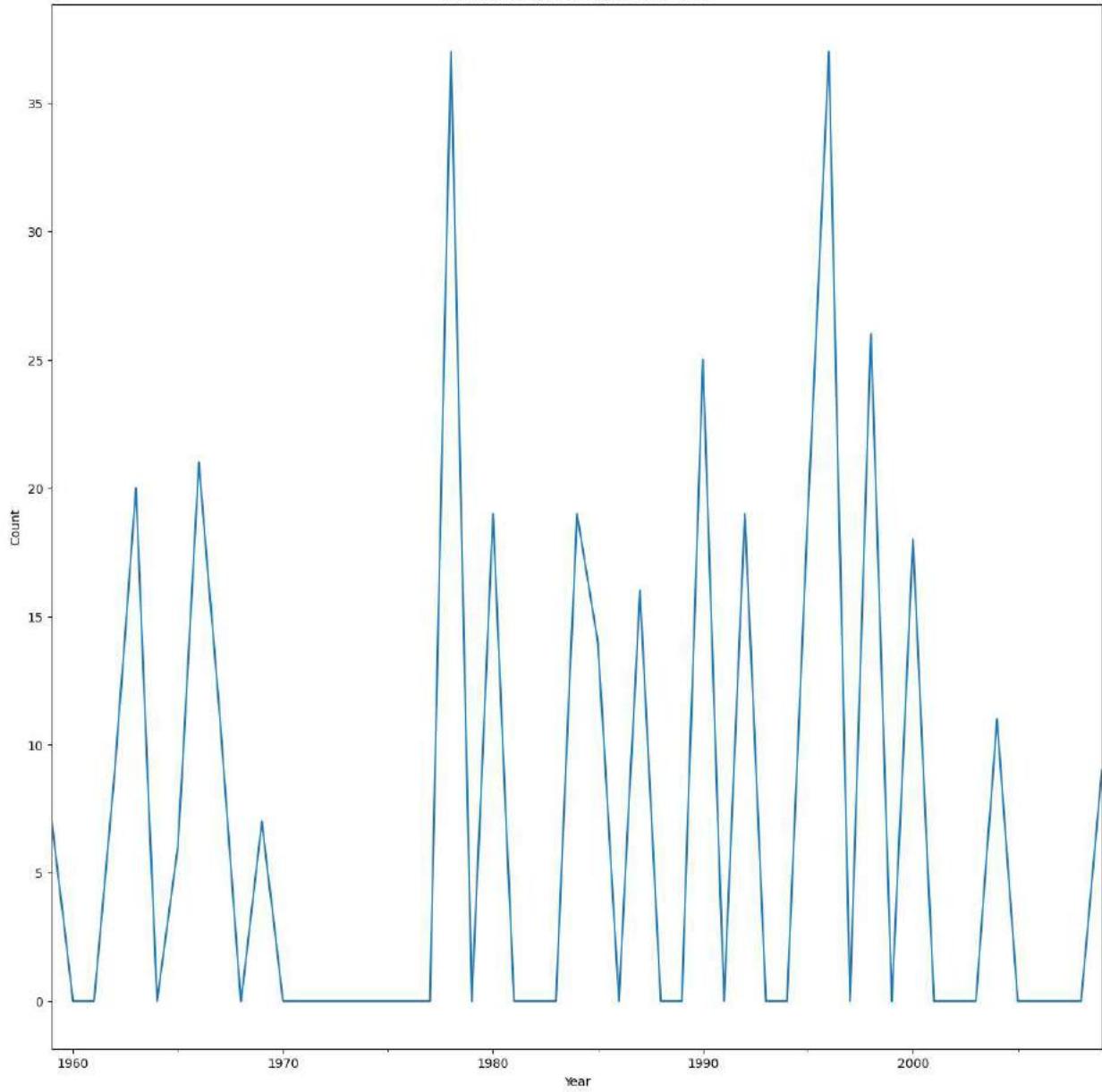
The code block for categorical data analysis and visualization counts the number of individuals in each group and creates a bar chart to visualize the distribution of individuals across groups. This can be useful for identifying patterns and trends in the data.

You could use this code block to analyze the number of astronauts from each country or the number of astronauts with each educational background.

```
In [80]: df['Year'] = pd.to_datetime(df['Year'], format='%Y')
df.set_index('Year', inplace=True)
```

```
In [94]: plt.figure(figsize=(15, 15))
df.resample('Y').size().plot()
plt.title('Number of Space Flights Over Time')
plt.xlabel('Year')
plt.ylabel('Count')
plt.show()
```

Number of Space Flights Over Time



Time Series Analysis and Visualization:

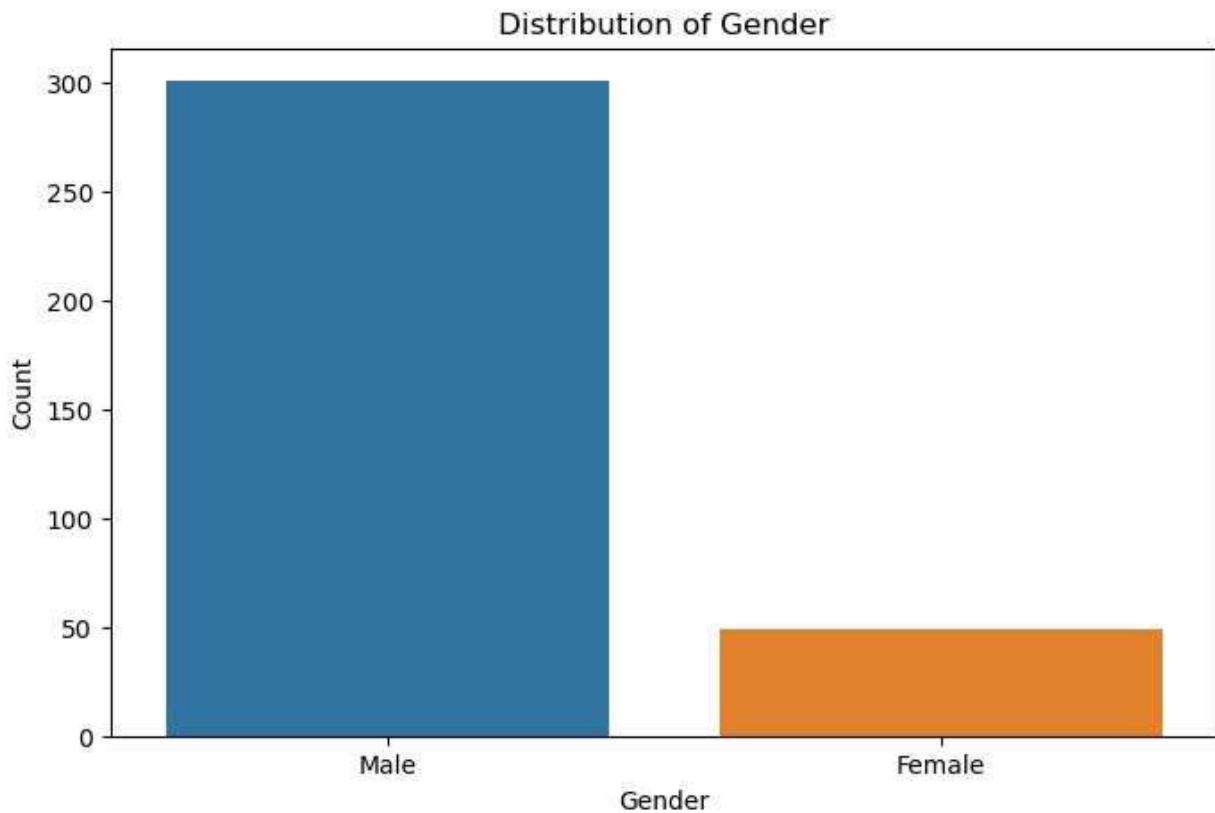
The code block for time series analysis and visualization converts the Year column to a datetime object and sets it as the index for time-based analysis. Then, it plots the number of space flights over time to observe trends.

This can be useful for identifying patterns and trends in the data, such as whether the number of space flights has increased or decreased over time.

You could use this code block to analyze the number of space flights launched by NASA each year or the number of space flights launched by each country each year.

```
In [82]: plt.figure(figsize=(8, 5))
sns.countplot(data=df, x='Gender')
plt.title('Distribution of Gender')
plt.xlabel('Gender')
```

```
plt.ylabel('Count')
plt.show()
```



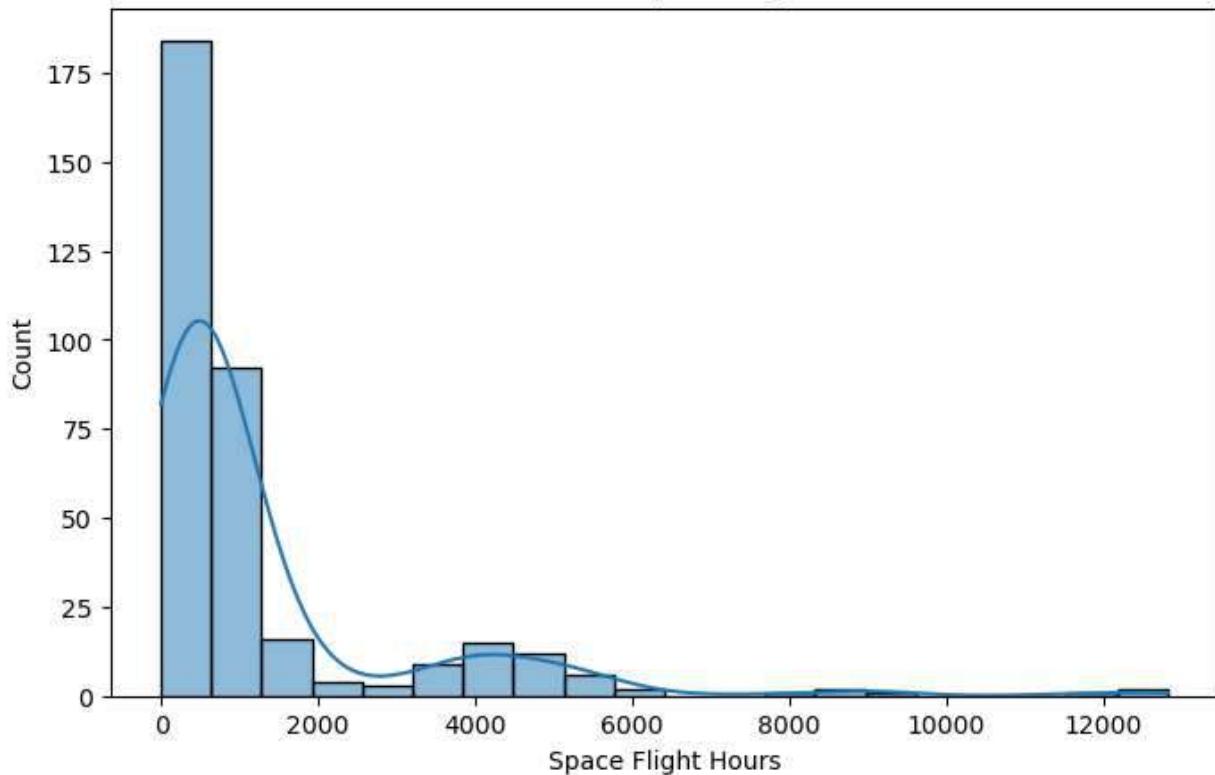
Gender Analysis:

The code block for gender analysis visualizes the distribution of gender in the dataset. This can be useful for identifying any gender imbalances in the data.

You could use this code block to analyze the distribution of gender among astronauts or the distribution of gender among scientists at a particular research institution.

```
In [83]: plt.figure(figsize=(8, 5))
sns.histplot(df['Space_Flight_hr'], bins=20, kde=True)
plt.title('Distribution of Space Flight Hours')
plt.xlabel('Space Flight Hours')
plt.ylabel('Count')
plt.show()
```

Distribution of Space Flight Hours



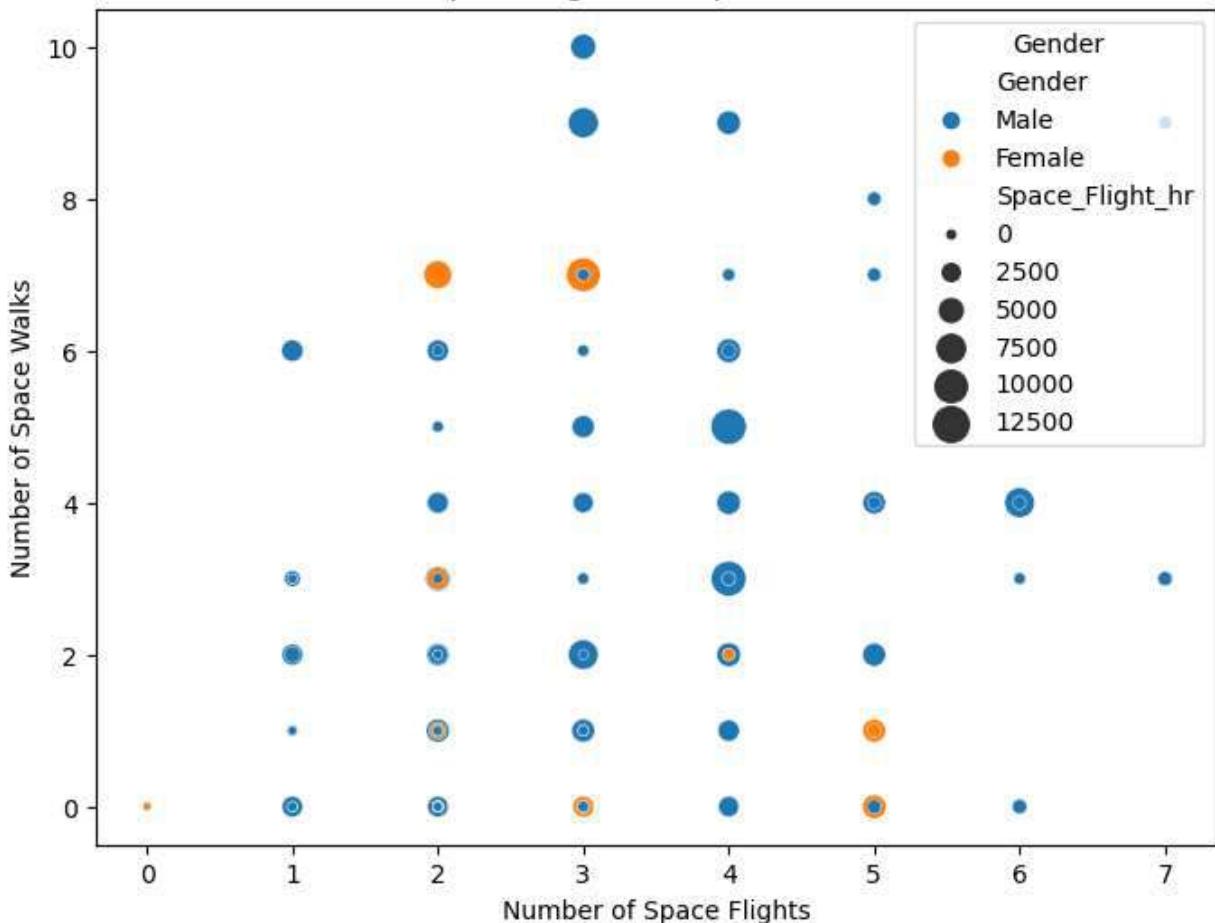
Numerical Data Analysis and Visualization:

The code block for numerical data analysis and visualization explores the distribution of space flight hours using a histogram. Then, it analyzes the relationship between space flights and spacewalks using a scatter plot and visualizes the distribution of space flight hours by gender using a box plot.

Histograms can be used to identify the shape of a distribution, such as whether it is normal or skewed.

```
In [84]: plt.figure(figsize=(8, 6))
sns.scatterplot(data=df, x='Space_Flights', y='Space_Walks', hue='Gender', size='Space
plt.title('Space Flights vs. Space Walks')
plt.xlabel('Number of Space Flights')
plt.ylabel('Number of Space Walks')
plt.legend(title='Gender', loc='upper right')
plt.show()
```

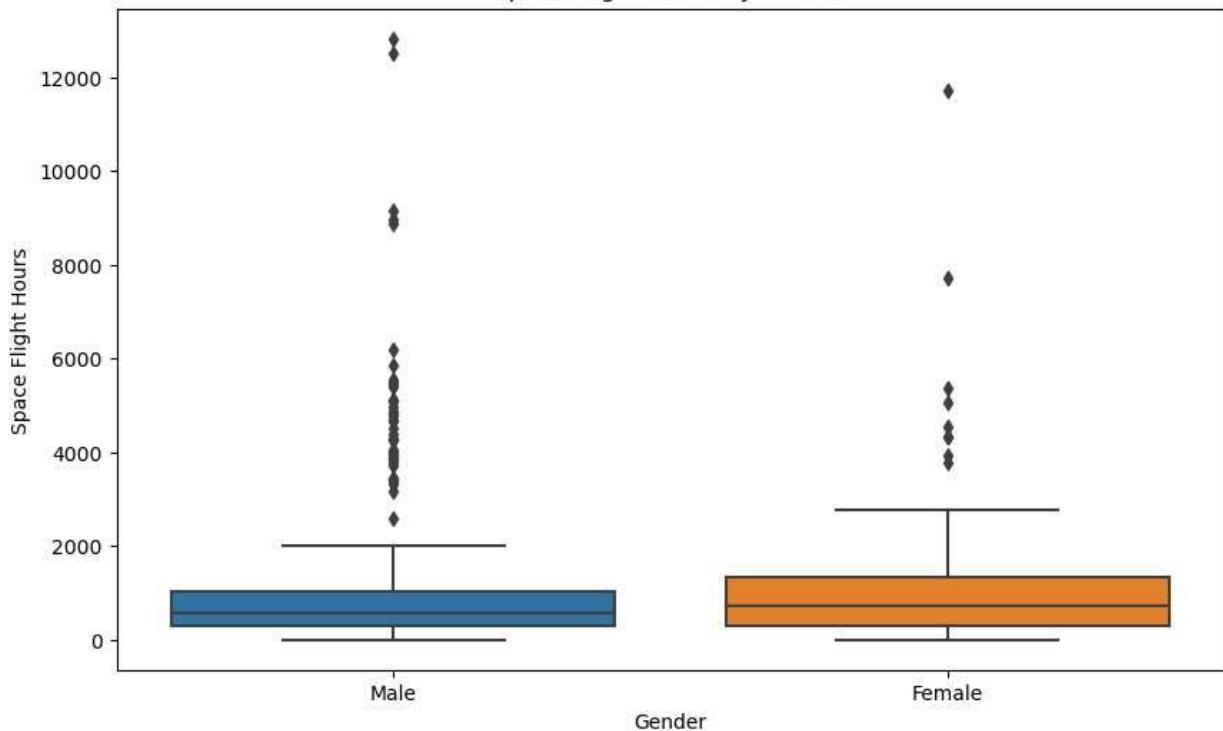
Space Flights vs. Space Walks



Scatter plots can be used to identify relationships between two variables, such as whether there is a correlation between space flights and spacewalks.

```
In [85]: plt.figure(figsize=(10, 6))
sns.boxplot(data=df, x='Gender', y='Space_Flight_hr')
plt.title('Space Flight Hours by Gender')
plt.xlabel('Gender')
plt.ylabel('Space Flight Hours')
plt.show()
```

Space Flight Hours by Gender



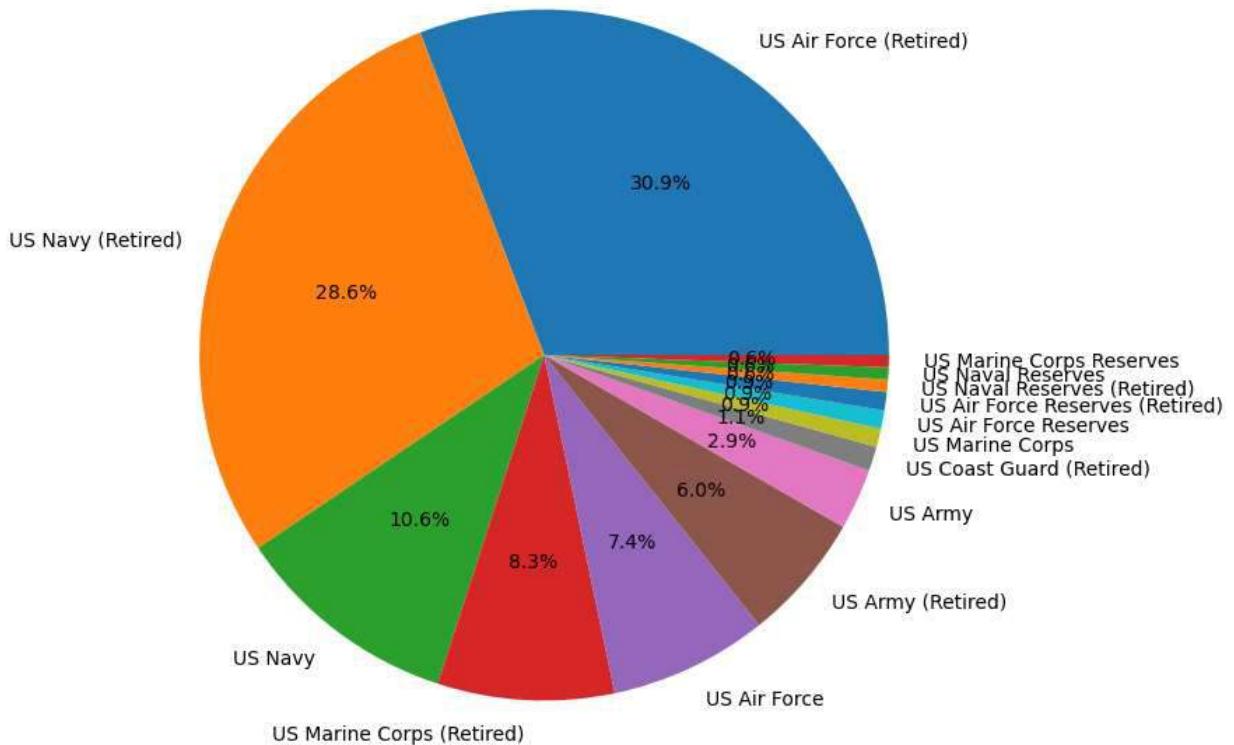
Box plots can be used to compare the distribution of a variable across different groups, such as the distribution of space flight hours by gender.

You could use this code block to analyze the distribution of space flight hours among astronauts, the relationship between space flights and spacewalks, or the distribution of space flight hours by gender among astronauts.

```
In [86]: military_counts = df['Military_Branch'].value_counts()
```

```
In [87]: plt.figure(figsize=(8, 8))
plt.pie(military_counts, labels=military_counts.index, autopct='%1.1f%%')
plt.title('Distribution of Individuals by Military Branch')
plt.show()
```

Distribution of Individuals by Military Branch



Military Branch Analysis and Visualization:

The code block for military branch analysis and visualization analyzes the distribution of individuals across military branches and creates a pie chart to visualize the distribution of individuals in each military branch.

This can be useful for identifying patterns and trends in the data, such as which military branch is the most common among astronauts.

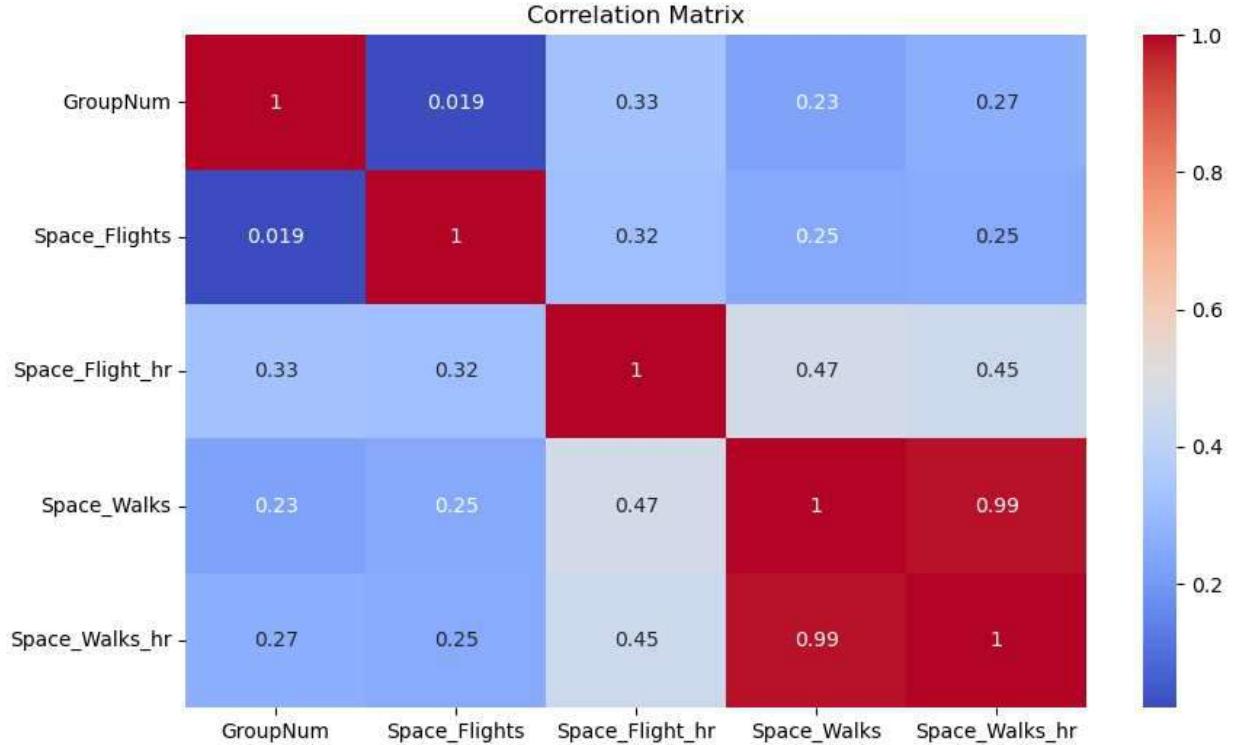
You could use this code block to analyze the distribution of military branches among astronauts or the distribution of military branches among scientists at a particular research institution.

In [88]:

```
correlation_matrix = df.corr()
plt.figure(figsize=(10, 6))
sns.heatmap(correlation_matrix, annot=True, cmap='coolwarm')
plt.title('Correlation Matrix')
plt.show()
```

C:\Users\hp\AppData\Local\Temp\ipykernel_8916\2935713837.py:1: FutureWarning: The default value of numeric_only in DataFrame.corr is deprecated. In a future version, it will default to False. Select only valid columns or specify the value of numeric_only to silence this warning.

```
correlation_matrix = df.corr()
```



Correlation Analysis:

The code block for correlation analysis calculates and visualizes the correlation between numerical columns in the dataset.

Correlation analysis can be used to identify relationships between variables, such as whether there is a correlation between age and experience or between education level and salary.

You could use this code block to analyze the correlation between space flight hours and experience, the correlation between education level and salary, or the correlation between age and health.