Analysis Report

MG

Modified: 2023-03-07

Contents

1 Business Task Statement

The purpose of this study is to get some insights of space missions that took place from 1957-2022. The data for this study is located at [1] and contains 4,630 records regarding space missions around the world for the period 10/4/1957 - 07/29/2022.

More specifically, this report finds answers to the following questions
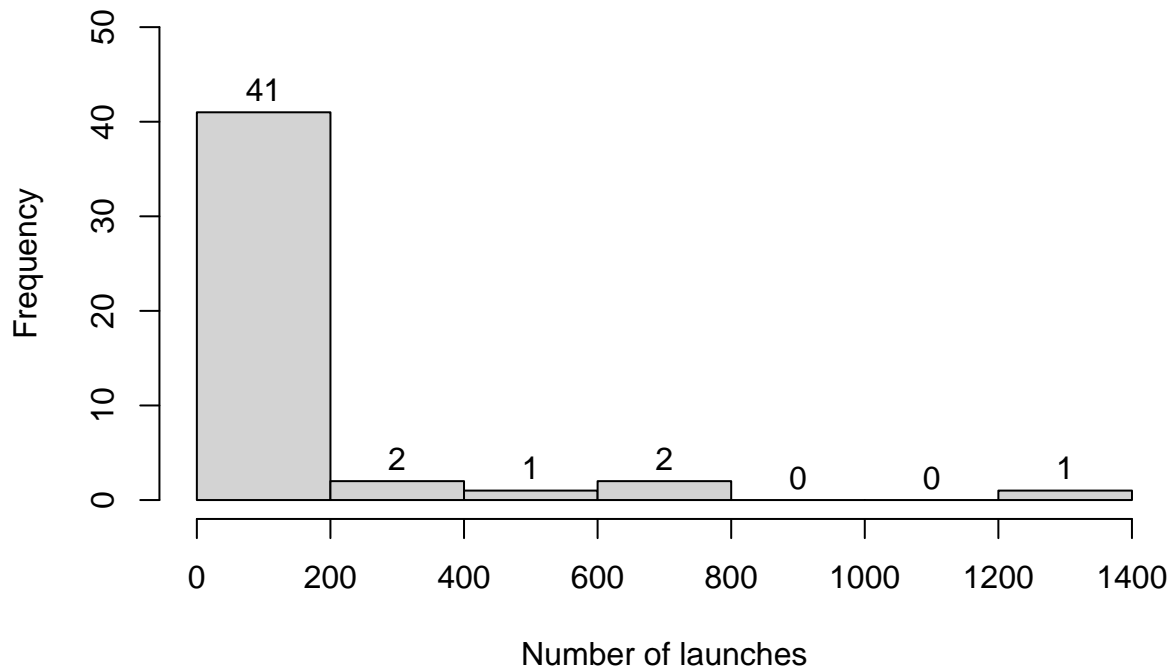
- Launches:
    - How many launches in total
    - How many successful/failed launches?
- Facilities:
    - How many launches per facility?
    - Which countries launched space missions?
- Rockets:
    - How many rockets?
    - How many is active vs. retiered?
    - Which rockets were used the most?
    - Which rockets served the longest?

2 Findings

Findings regard period 10/4/1957 - 7/29/2022, unless stated otherwise.

- There were 4630 launches in total.
- There were 47 facilities capable of rocket launch.
- Facilities were located in 20 countries (or areas).
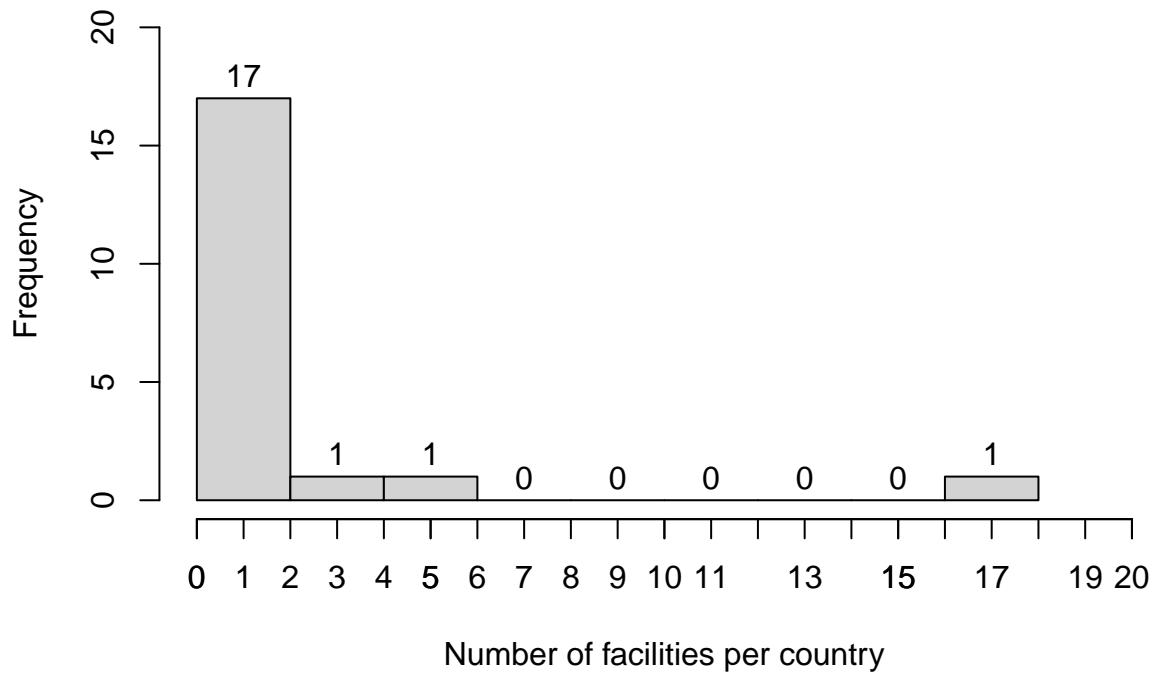- 87% of facilities performed no more than 200 launches in total.

Histogram of the total number of launches by facilities all over the world for the period 10/4/1957 – 7/29/2022.

2.1 Facilities

- A country can have several launching facilities.

Histogram of the number of facilities per country for the period 10/4/1957 – 7/29/2022.

Frequency

Number of facilities per country

- Facilities were located in 20 countries.
- There were two locations with undefined country of ownership: Yellow Sea, and Barents Sea, where rockets were launched from submarines (Barents Sea) or barges (Yellow Sea).

Table 1: Facilities per location/country for the period 10/4/1957 - 7/29/2022.

Area	FacilityCount	Percent
USA	17	36
Russia	5	11
China	4	9
Iran	2	4
Japan	2	4
North Korea	2	4
Yellow Sea	2	4
Algeria	1	2
Australia	1	2
Barents Sea	1	2
Brazil	1	2
France	1	2
Gran Canaria	1	2
India	1	2
Israel	1	2
Kazakhstan	1	2
Kenya	1	2
Kiribati	1	2
New Zealand	1	2
South Korea	1	2

Percent of facilities located in a country for the period 10/4/1957 – 7/29/2022.

Number of facilities: 47.

4
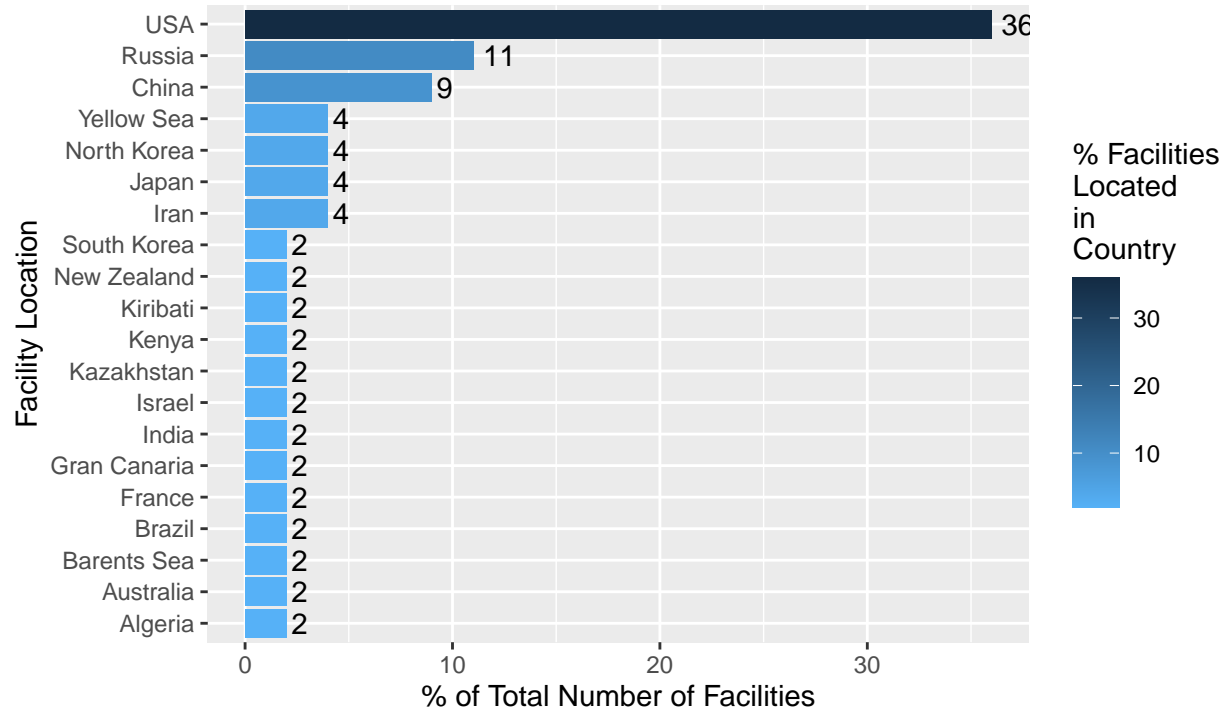
- Eight countries were used by more than one company. Top 5 most popular launching locations:
  - USA (used by 21 companies),
  - Kazakhstan (used by 11 companies),
  - Russia (used by 9 companies),
  - China (used by 8 companies),
  - Japan (used by 4 companies).

The number of companies launching in a given country (10/4/1957 – 7/29/2022).

Yellow Sea and Barents Sea – launches from barges and submarines.
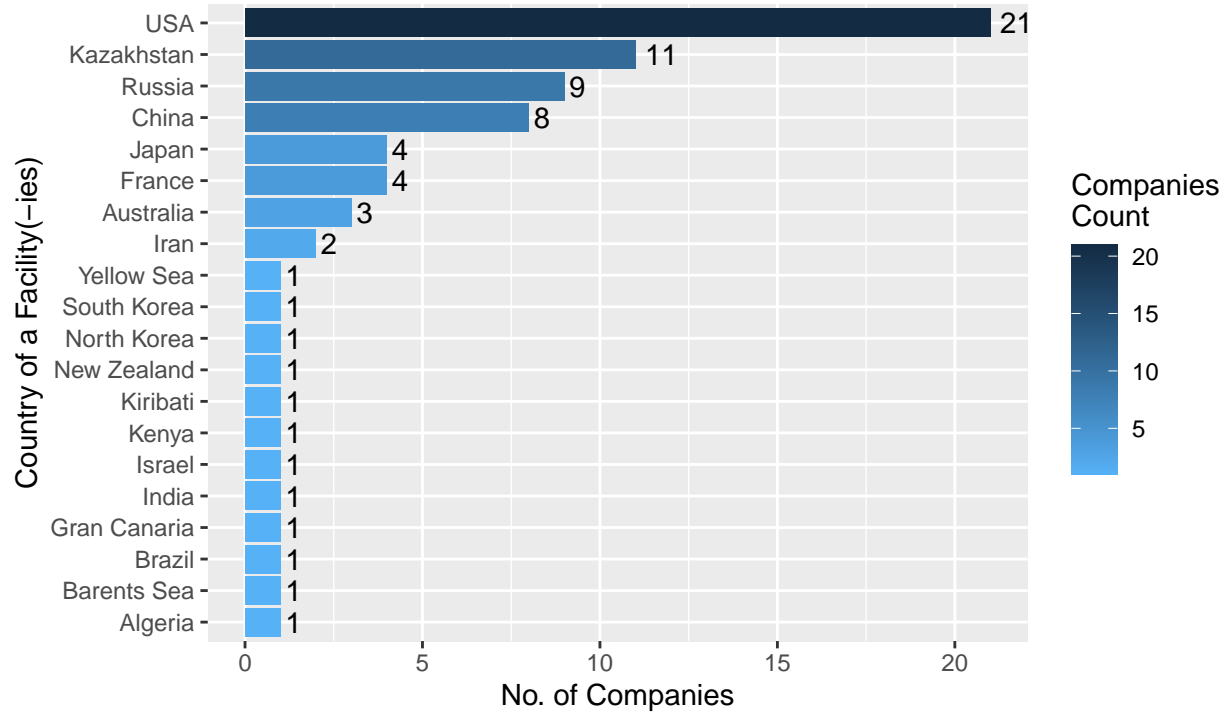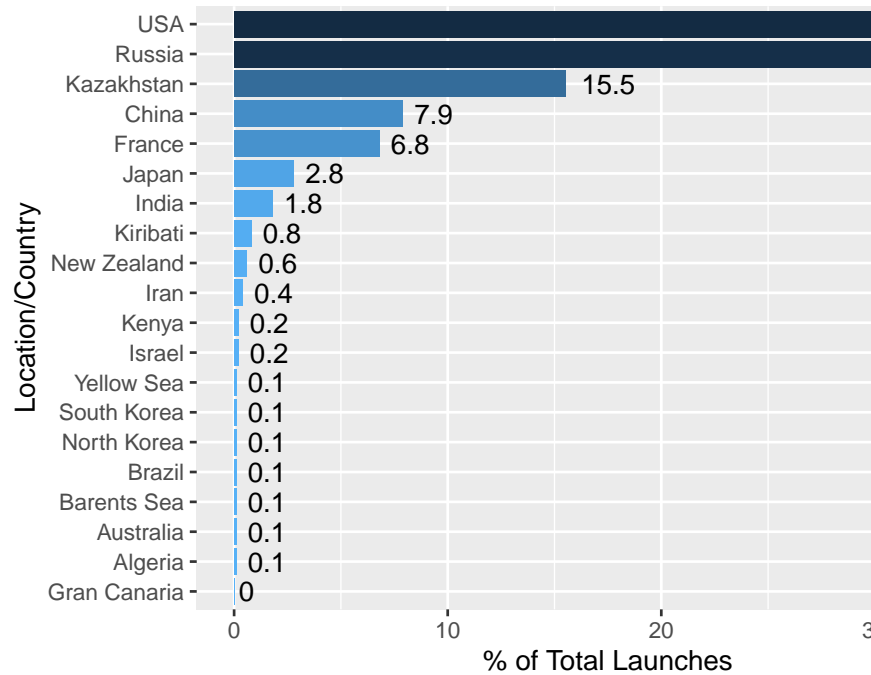

2.2 Companies

- There were 62 companies that launched space missions.
- 50 companies launched in a single country/location.
- There were twelve companies that launched space missions in two different countries.

2.3 Launches

- Top 5 countries/locations (USA, Russia, Kazahstan, China, France) account for over 92% of global

Locations/countries w.r.t. the number of launch
the period 10/4/1957 – 7/29/2022.

There were 4630 launches in total.

Location/Country	% of Total Launches
USA	
Russia	
Kazakhstan	15.5
China	7.9
France	6.8
Japan	2.8
India	1.8
Kiribati	0.8
New Zealand	0.6
Iran	0.4
Kenya	0.2
Israel	0.2
Yellow Sea	0.1
South Korea	0.1
North Korea	0.1
Brazil	0.1
Barents Sea	0.1
Australia	0.1
Algeria	0.1
Gran Canaria	0

launches. There were 4630 launches in total.

- Top 10 launching facilities taking into account the number of launches.

Table 2: Locations/countries launching space missions ordered by percentage of the global number of launches for the period 10/4/1957 - 7/29/2022.

Area	Total	Percent
USA	1472	31.8
Russia	1416	30.6
Kazakhstan	719	15.5
China	365	7.9
France	314	6.8
Japan	130	2.8
India	82	1.8
Kiribati	36	0.8
New Zealand	28	0.6
Iran	17	0.4
Israel	11	0.2
Kenya	9	0.2
Australia	6	0.1
North Korea	5	0.1
South Korea	5	0.1
Algeria	4	0.1
Barents Sea	3	0.1
Brazil	3	0.1
Yellow Sea	3	0.1
Gran Canaria	2	0.0

Top 10 facilities w.r.t. the total number of launches
for the period 10/4/1957 – 7/29/2022.
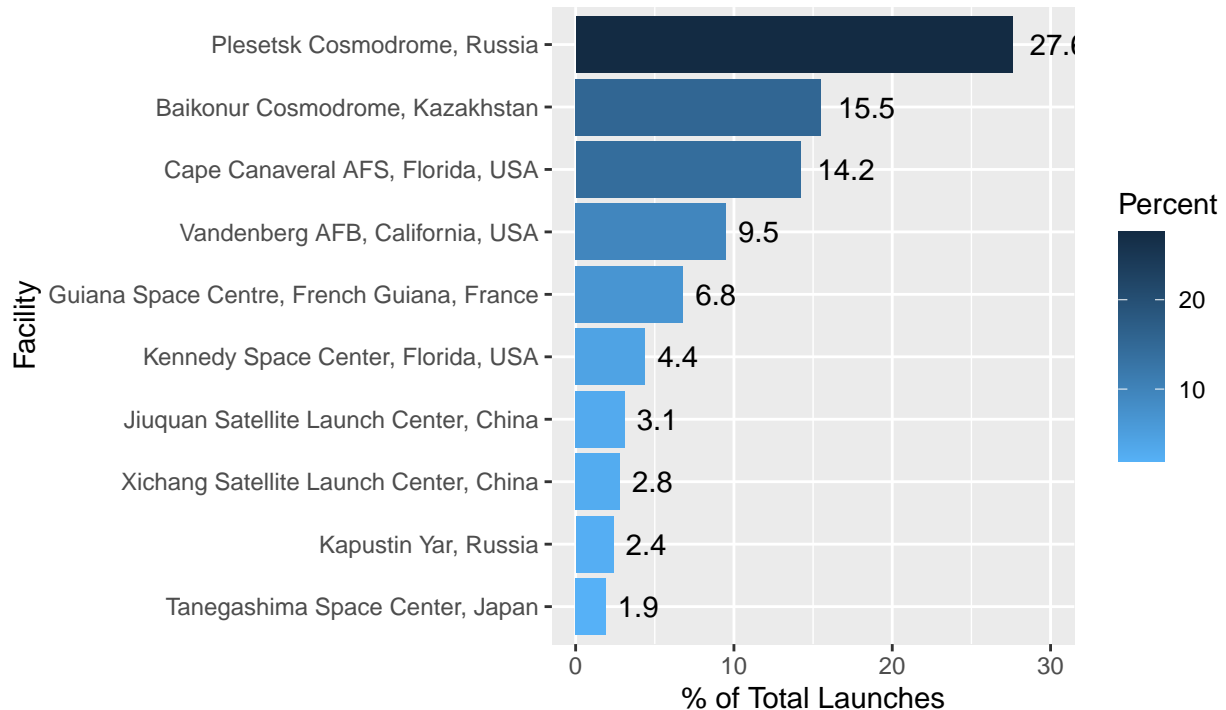
There were 4630 launches in total.

Table 3: The number of successful and failed launches per facility for the period 10/4/1957 - 7/29/2022.

	x
Total	4630
Success	4162
Failure	357
PartialFailure	107
PrelaunchFailure	4

- 90% (4162) of 4630 launches in total were successful.

**Space missions launch successes and failures
during 10/4/1957 – 7/29/2022. There were 4630 launches in total.**

–>
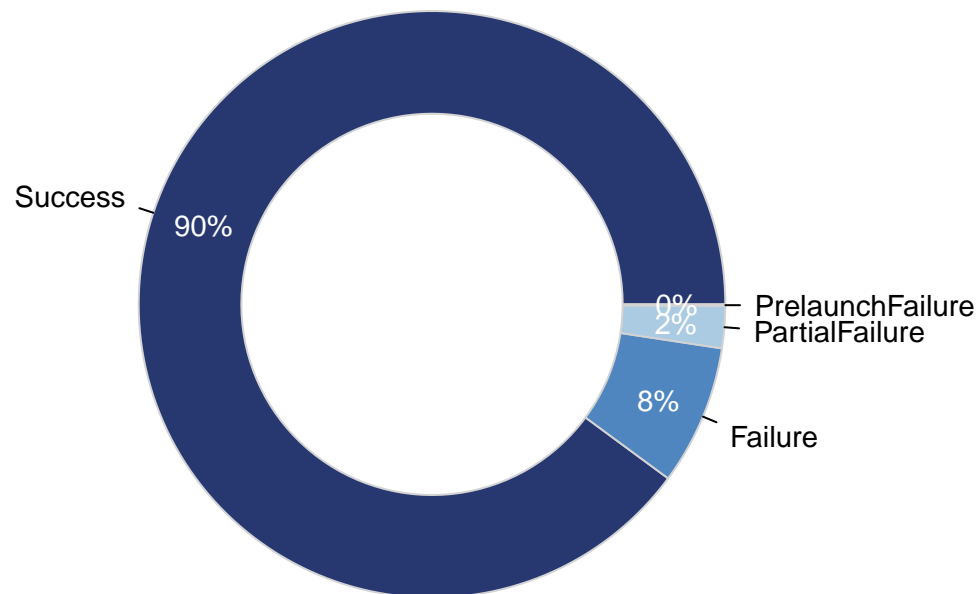
- Percentage of successful and failed launches is shown below:

Table 4: Top 10 launching facilities for the period 10/4/1957 - 7/29/2022.

Place	TotalPercent	Total	RelSuccessPerc	RelFailurePerc
Plesetsk Cosmodrome, Russia	27.6	1278	94.1	5.9
Baikonur Cosmodrome, Kazakhstan	15.5	719	86.9	13.1
Cape Canaveral AFS, Florida, USA	14.2	658	87.7	12.3
Vandenberg AFB, California, USA	9.5	442	87.3	12.7
Guiana Space Centre, French Guiana, France	6.8	314	94.3	5.7
Kennedy Space Center, Florida, USA	4.4	205	98.5	1.5
Jiuquan Satellite Launch Center, China	3.1	144	90.3	9.7
Xichang Satellite Launch Center, China	2.8	130	93.1	6.9
Kapustin Yar, Russia	2.4	112	85.7	14.3
Tanegashima Space Center, Japan	1.9	88	95.5	4.5

Percentage of successful and failed launches per facility for the period 10/4/1957 – 7/29/2022.
There were 4630 launches in total.
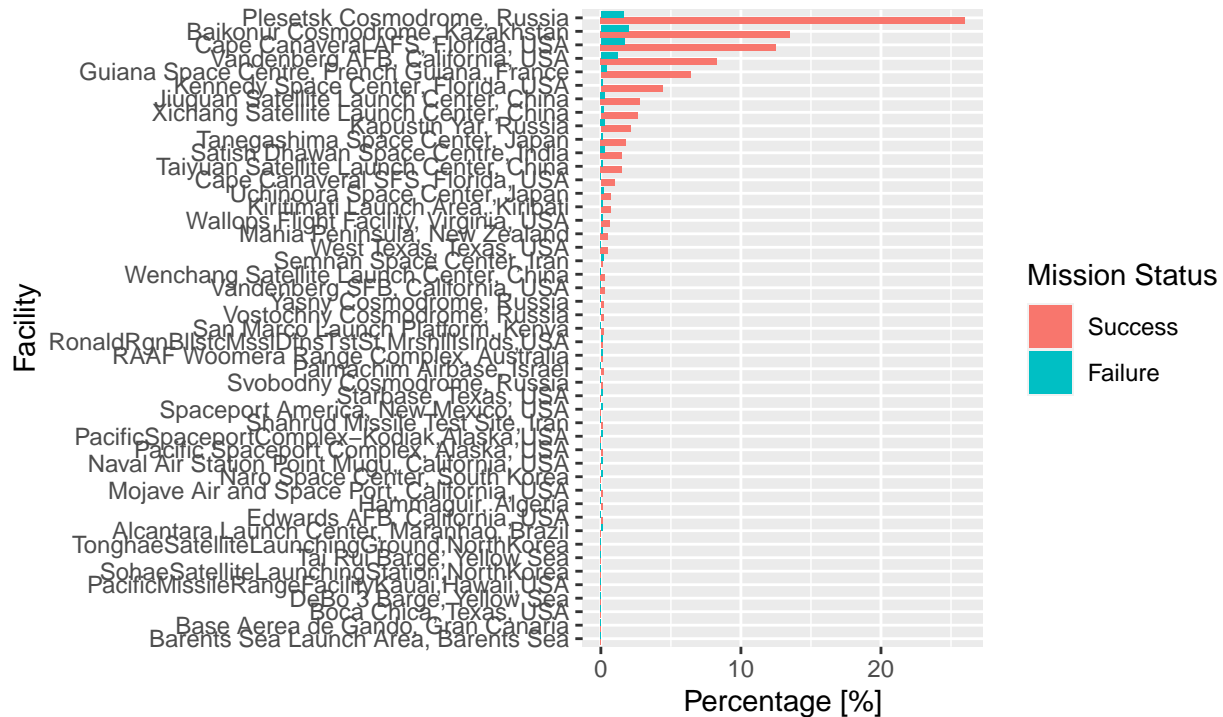
- Top 10 launching facilities with the highest number of space mission launches.

- Top 10 facilities with the highest number of launches ordered by successful launches relative to the total number of launches a given facility performed. Kennedy Space Center had the highest relative success value, however, it performed only 4.4% of all global launches.

Table 5: Facilities with most launches globally ordered by successful launches relative to the number of launches a given facility performed for the period 10/4/1957 - 7/29/2022.

Place	RelSuccessPerc	TotalPercent	Total
Kennedy Space Center, Florida, USA	98.5	4.4	205
Tanegashima Space Center, Japan	95.5	1.9	88
Guiana Space Centre, French Guiana, France	94.3	6.8	314
Plesetsk Cosmodrome, Russia	94.1	27.6	1278
Xichang Satellite Launch Center, China	93.1	2.8	130
Jiuquan Satellite Launch Center, China	90.3	3.1	144
Cape Canaveral AFS, Florida, USA	87.7	14.2	658
Vandenberg AFB, California, USA	87.3	9.5	442
Baikonur Cosmodrome, Kazakhstan	86.9	15.5	719
Kapustin Yar, Russia	85.7	2.4	112

Table 6: Stats for the launches for the period 10/4/1957 - 7/29/2022.

|	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
LaunchesCount	3	50	61.5	70.2	96	157

Top 10 most launching facilities
ordered by percentage of successful launches
relative to the total number of launches
a given facility performed
for the period 10/4/1957 – 7/29/2022.

The color represents percentage of launches
a given facility performed w.r.t. all launches
performed globally, 4630 in total.
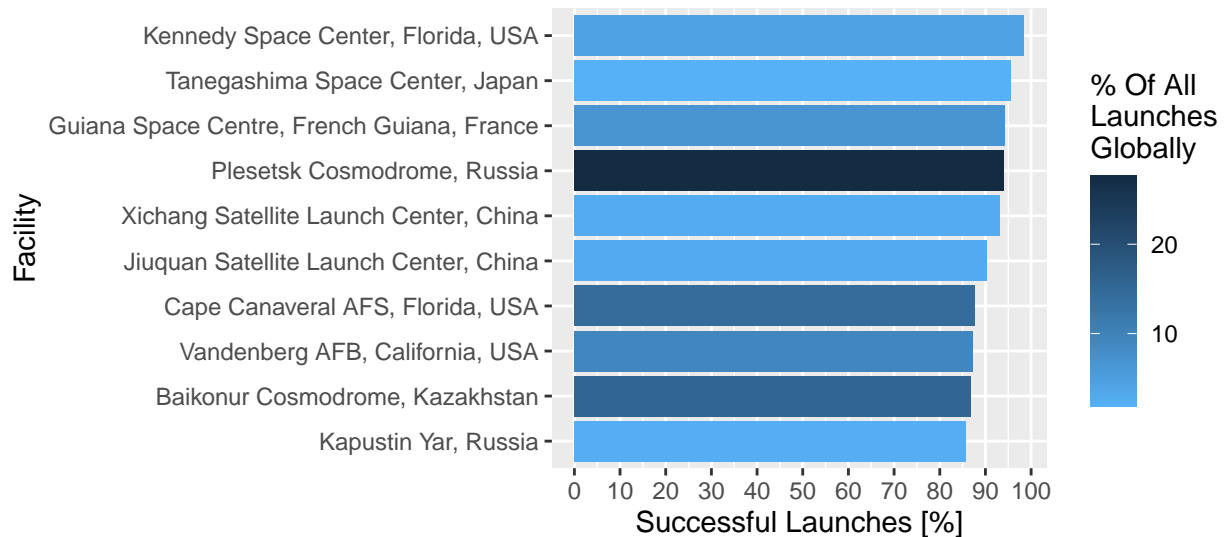
- Descriptive stats regarding launches w.r.t. launches per year for the period 10/4/1957 - 7/29/2022.:

- The number of launches was the highest in 1966-1978 (around 100 per year), and 2016-2022 (around 100 per year). The lowest number of launches per year was performed in 2001-2015 that ranged from 0.8% of launches counted globally (37) to 1.1% (53).
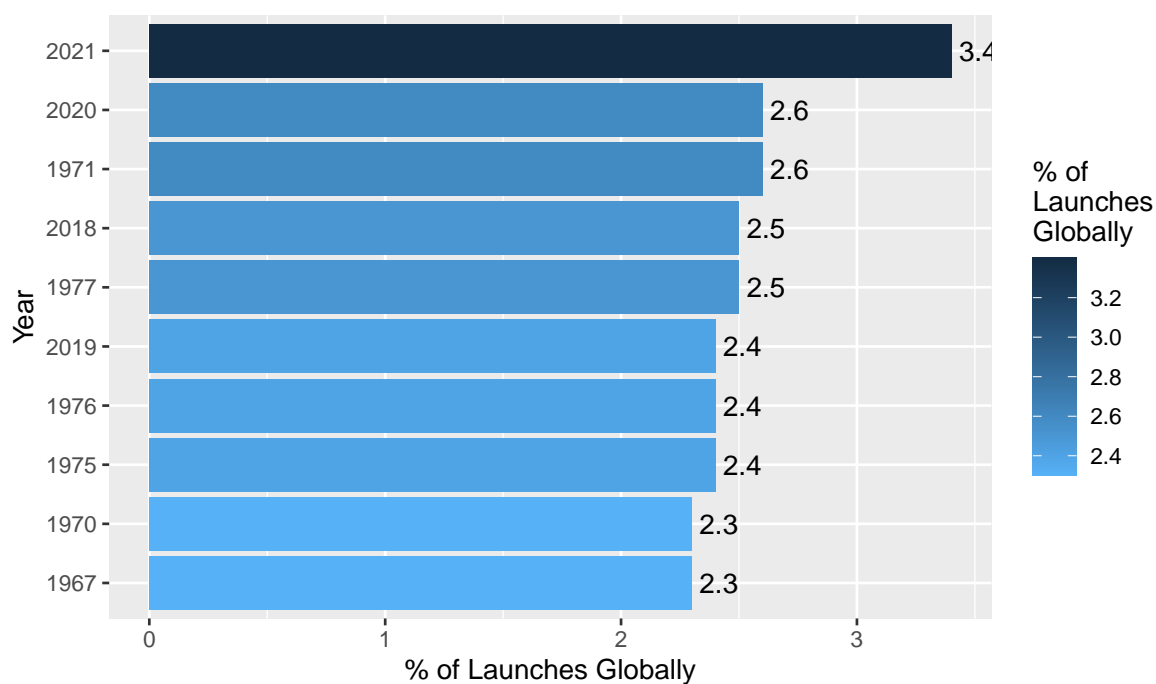
Percentage of launches per year globally for the period 10/4/1957 – 7/29/202

There were 4630 launches in total.
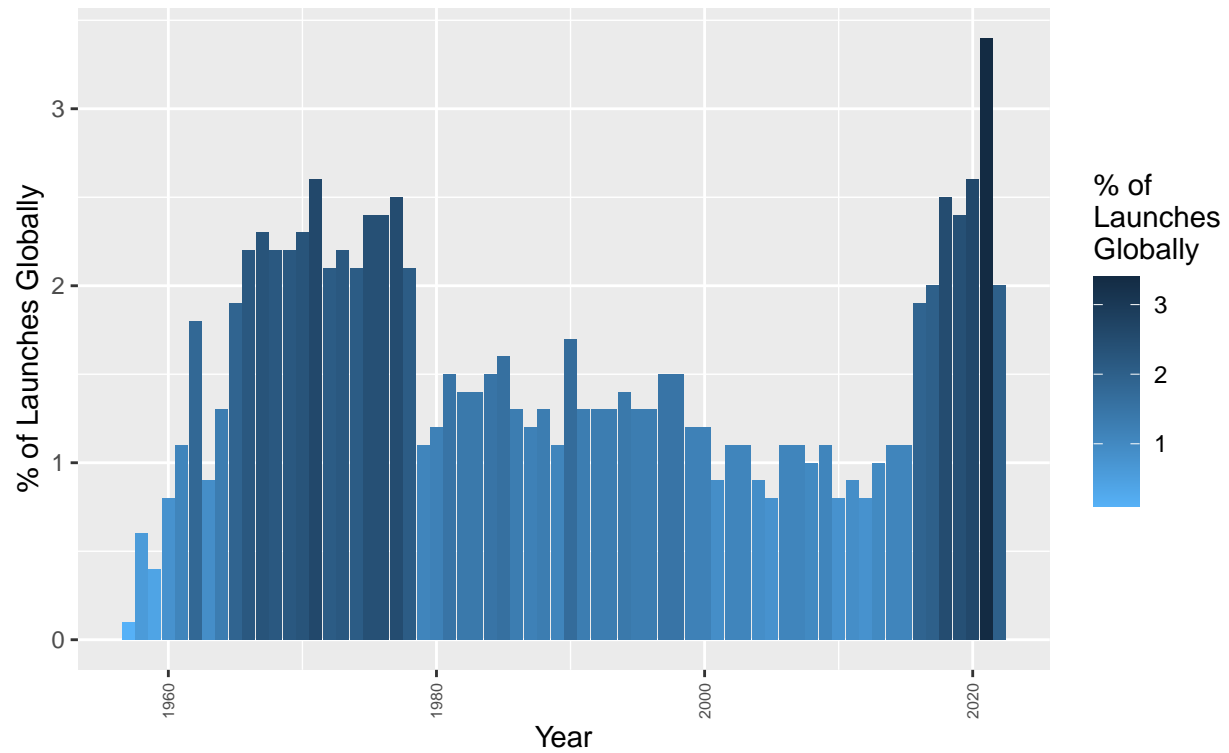
- The highest number of launches was in 2021 (157) and it accounted for 3.4% of global launches.

Top 10 years with the most launches a year for the period
 10/4/1957 – 7/29/2022.

There were 4630 launches in total.

- Year 1957 observed the lowest number of launches, i.e., 3 accounting for 0.1% of total launches for the period. However, 1957 started in April. The second lowest yearly number of launches was in 1959, i.e., 20 (0.4% of all global launches).

Bottom 10 years with the least launches a year for the period
 10/4/1957 – 7/29/2022.
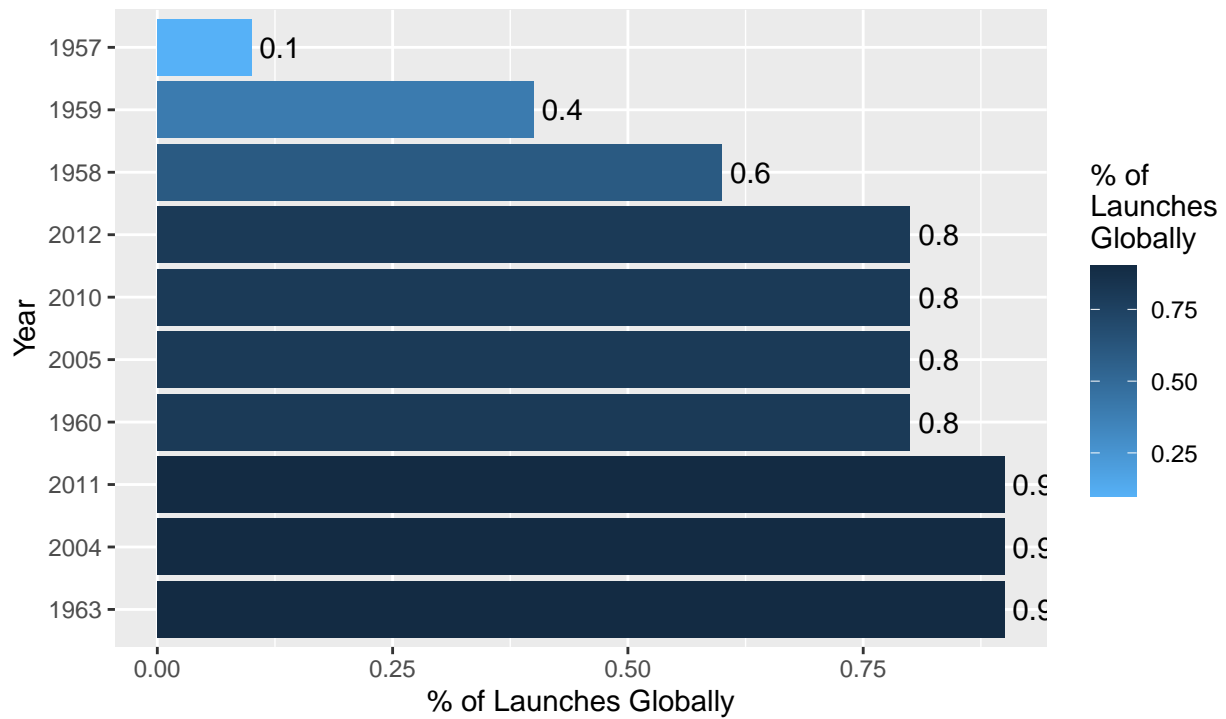
There were 4630 launches in total.

Table 7: Stats for rockets for the period 10/4/1957 - 7/29/2022.Total number of rockets used is 371

	LaunchCount	FirstLaunch	LastLaunch	OperationInYears
Min	Min. : 1.00	Min. :1957-10-04	Min. :1957-11-03	Min. : 0.000
Median	Median : 4.00	Median :1994-03-13	Median :2003-01-16	Median : 3.000
Mean	Mean : 12.48	Mean :1992-03-24	Mean :1997-12-25	Mean : 5.758
Max	Max. :446.00	Max. :2022-07-27	Max. :2022-07-29	Max. :43.800

2.4 Rockets

- There were total 371 rockets used for launching. Most rockets (97%) performed no more than 50 launches.

Histogram of number of launches per rocket for the period 10/4/1957 – 7/29/2022. Total number of rockets: 371 Total number of launches: 4630.
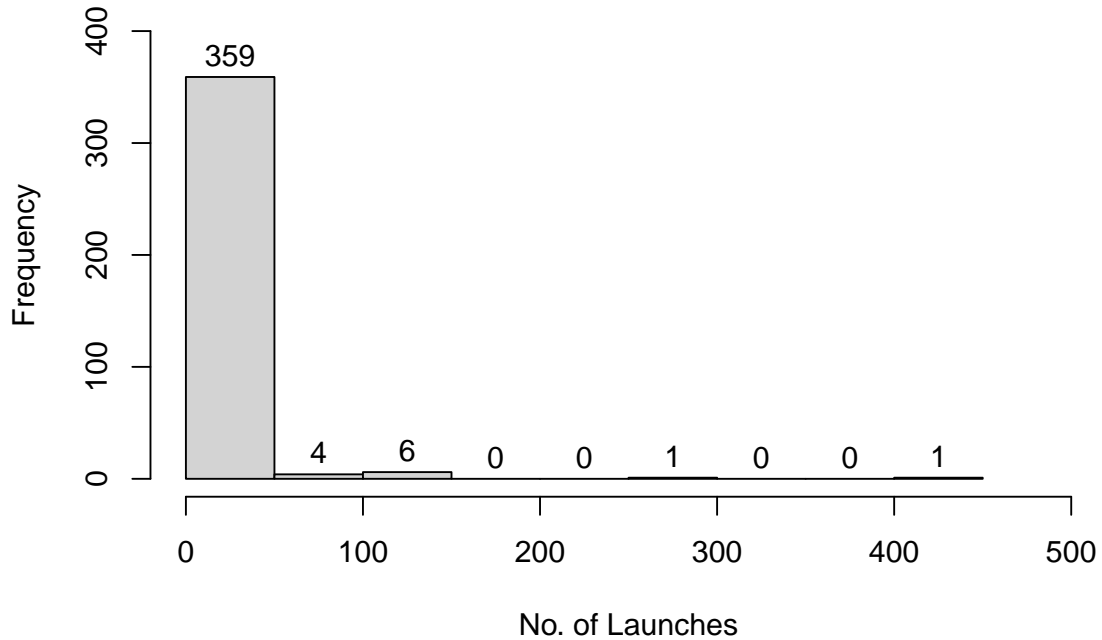
- The status of rockets: active vs. retired.

Active vs. retired rockets as on 7/29/2022 for the period 10/4/1957 – 7/29/2022.

Active

28%

Total number
of rockets = 371

72%

Retired
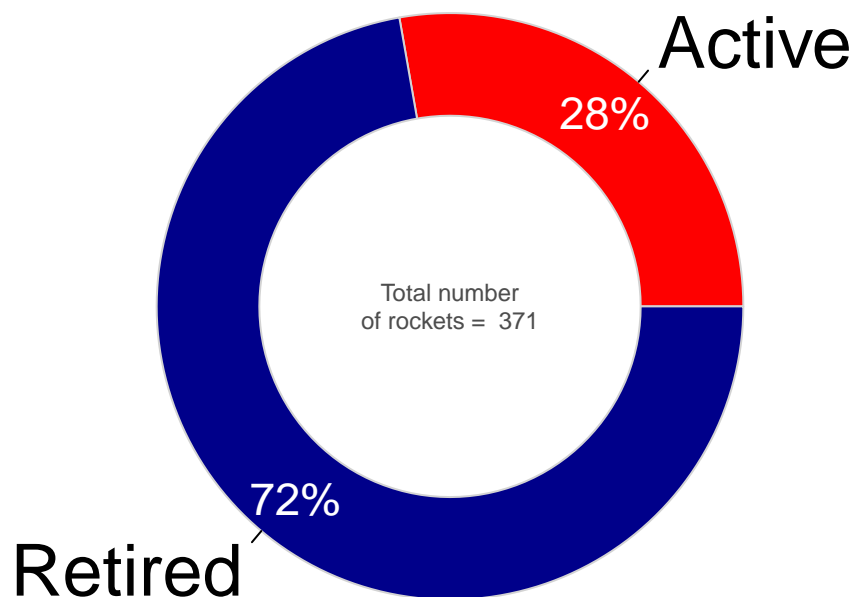
- Rocket Cosmos-3M (11K65M) was the most popular one with 446 launches which accounted for 9.6 of total launches.

Table 8: Most used rockets of all times for the period 10/4/1957 - 7/29/2022.Total number of rockets was 371.

Rocket	LaunchCount	CountPerc	RocketStatusName	FirstLaunch	LastLaunch	OperationInYea
Cosmos-3M (11K65M)	446	9.6	Retired	1967-05-15	2010-04-27	4
Voskhod	299	6.5	Retired	1963-11-16	1976-06-29	1
Molniya-M /Block ML	128	2.8	Retired	1974-04-20	2005-06-21	3
Cosmos-2I (63SM)	126	2.7	Retired	1965-10-19	1977-06-18	1
Soyuz U	125	2.7	Retired	1973-05-18	2017-02-22	4
Tsyklon-3	122	2.6	Retired	1977-06-24	2009-01-30	3
Falcon 9 Block 5	111	2.4	Active	2018-05-11	2022-07-24	4
Tsyklon-2	106	2.3	Retired	1969-08-06	2006-06-25	3
Vostok-2M	93	2.0	Retired	1964-08-28	1991-08-29	2
Molniya-M /Block 2BL	87	1.9	Retired	1972-09-19	2010-09-30	3

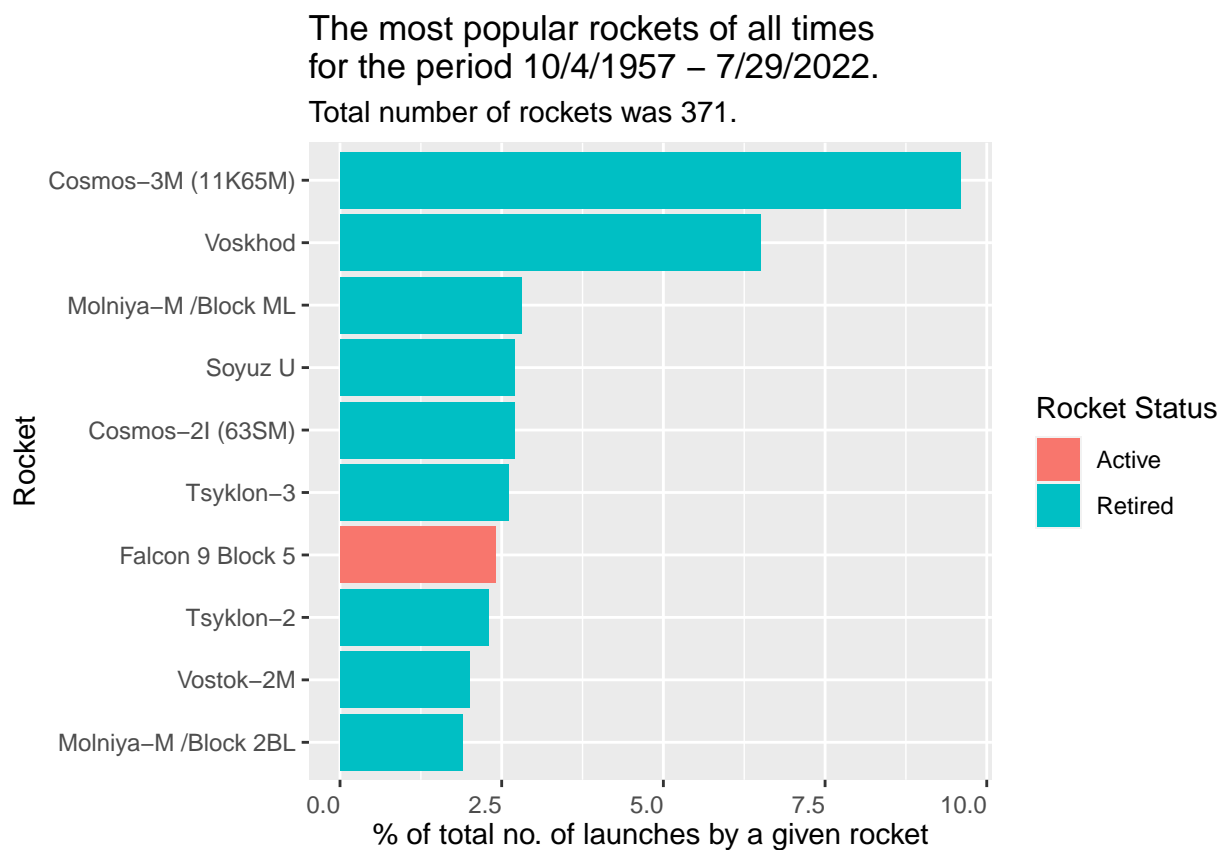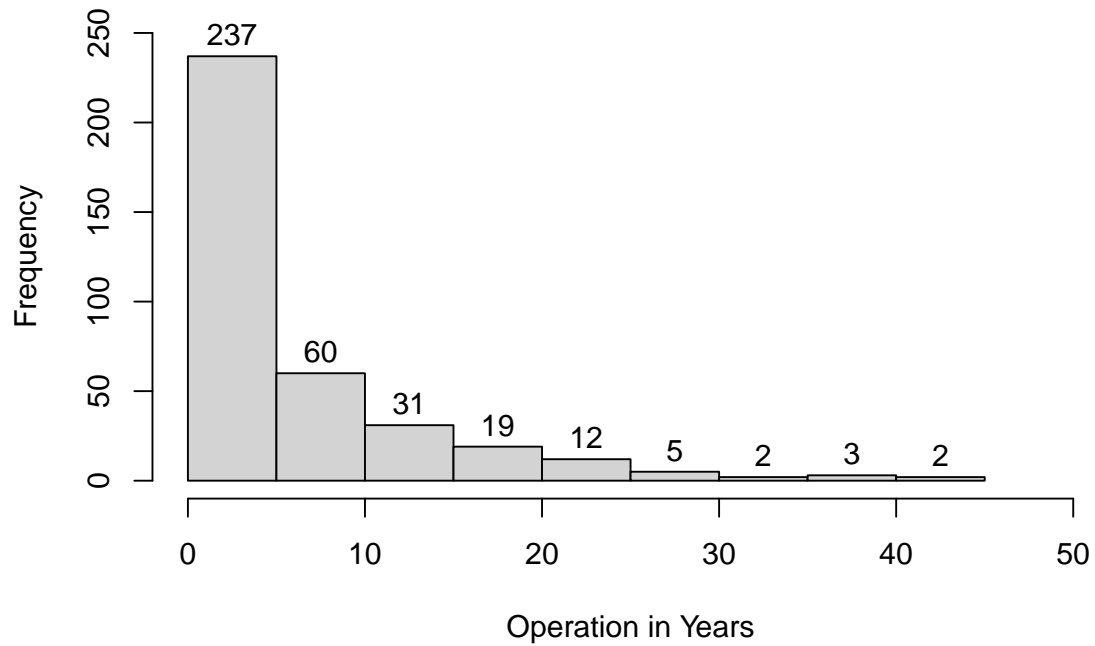The most popular rockets of all times for the period 10/4/1957 – 7/29/2022. Total number of rockets was 371.

- There were total 371 rockets used for launching. Most rockets (64%) operated for no more than 5 years.
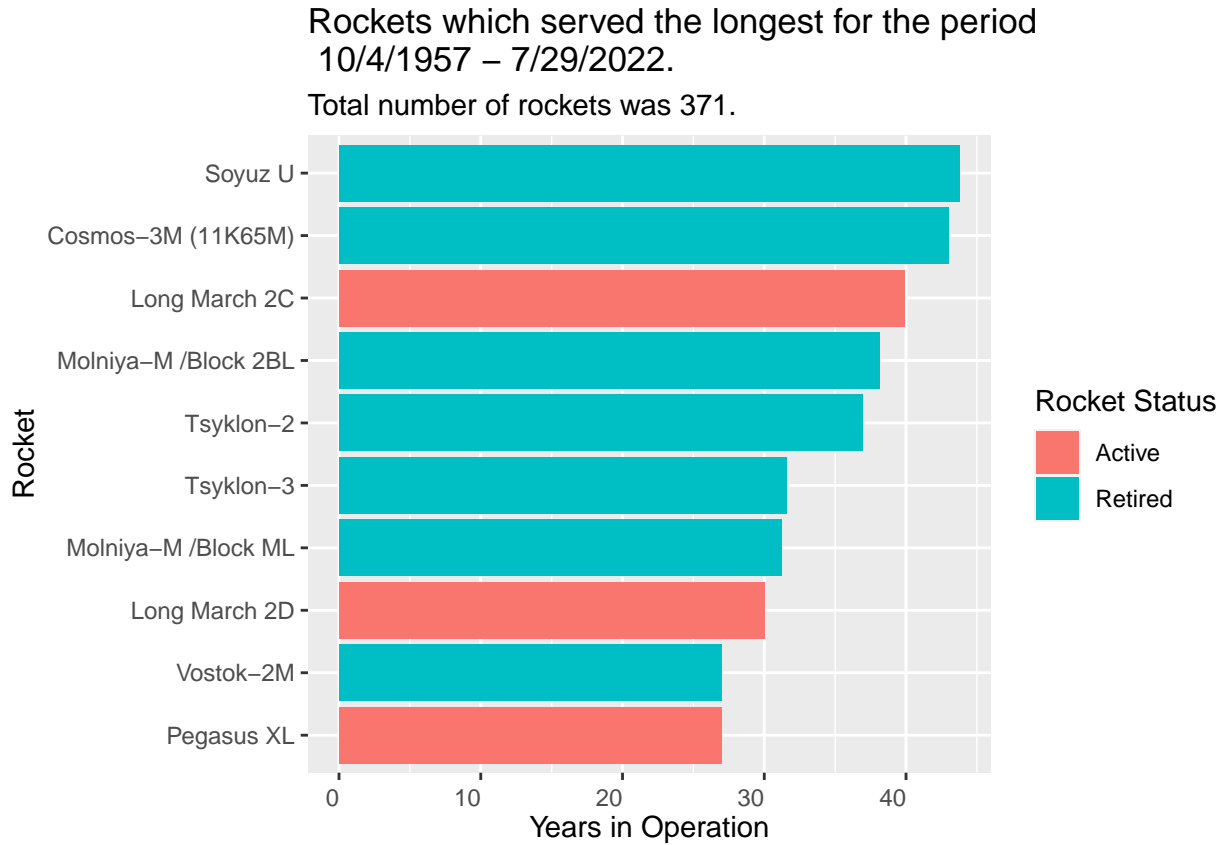
**Histogram of operational rocket lifespan for the period
10/4/1957 – 7/29/2022. Total number of rockets: 371.**

- The longest serving rocket was Soyuz U which served from 1233 and retired on 17219, accounting for 43.8 years of operation. The longest serving rocket still active as of the end of period for this dataset was Long March 2C with 39.9 years of operation.

Table 9: Rockets that served the longest for the period 10/4/1957 - 7/29/2022.Total number of rockets was 371.

Rocket	LaunchCount	CountPerc	RocketStatusName	FirstLaunch	LastLaunch	OperationInYea
Soyuz U	125	2.7	Retired	1973-05-18	2017-02-22	43
Cosmos-3M (11K65M)	446	9.6	Retired	1967-05-15	2010-04-27	43
Long March 2C	43	0.9	Active	1982-09-09	2022-07-15	39
Molniya-M /Block 2BL	87	1.9	Retired	1972-09-19	2010-09-30	38
Tsyklon-2	106	2.3	Retired	1969-08-06	2006-06-25	36
Tsyklon-3	122	2.6	Retired	1977-06-24	2009-01-30	31
Molniya-M /Block ML	128	2.8	Retired	1974-04-20	2005-06-21	31
Long March 2D	44	1.0	Active	1992-08-09	2022-07-29	30
Vostok-2M	93	2.0	Retired	1964-08-28	1991-08-29	27
Pegasus XL	30	0.6	Active	1994-06-27	2021-06-13	27

Rockets which served the longest for the period 10/4/1957 – 7/29/2022.

Total number of rockets was 371.

3 Preparing Data

The original data is located at [1]. The data is organized in two .csv files:

- `space_missions.csv` - the actual data
- `space_missions_data_dictionary.csv` - description of fields

Data Structure	No. of Fields	No. of Records	Date Added
Single table	9	4,630 + header	08/03/2022

3.1 Data Structure

Each observation consists of 9 attributes that describe: Company, Location, Date, Time, Rocket, Mission, RocketStatus, Price, MissionStatus.

- Company - a text field that can includes spaces, variable length
- Location - a text field, variable length, includes: the site, place, state, country. State and country is for USA. Tokens separated by comma.
- Date - m/d/yyyy; m and d can be a single or double digit
- Time - h:m:ss; h and m can be a single or double digit; according description UTC
- Rocket - a text that can have multiple tokens
- Mission - a text that can contain multiple tokens; letter encoding issues
- RocketStatus - textual one word (Retired, Active)
- Price - numeric, often missing
- MissionStatus - (Success, Failure, Partial Failure, Prelaunch Failure)

3.2 Data Credibility

The data is provided by a respected website devoted to Data Analytics [1]. I checked and confirmed from other sources a couple randomly selected observations:

- Falcon 9 Block 5 | Nilesat-301 | SpaceX | June 8th, 2022
- Martin Marietta SLC-4E, Vandenberg AFB, California, USA 11/8/1991 7:07:00 Titan IV(403)A SLDCOM & NOSS 3x

3.3 Data Bias

To rule out the bias in the data, the research should be performed to verify if all space missions were indeed included within the specified timeline, i.e., 1957-2022. My assumptions is that the data is collected and presented in a fair manner.

3.4 License

I have not found any license on how to use the data. I have reached out to the staff and they said that "they [datasets] are free to use as you wish".

3.5 Potential Issues

- Some data is missing. Especially, the price field.
- Mission field has different encoding for letters.
- Location field probably needs to be separated.

3.6 Preparation Methodology

I used Microsoft Excel to visually inspect the data and address the issues.

4 Data Processing

I could not use pivot tables for the dataset due to a MS Excel memory issue: when trying to create a pivot table from the data csv file with 4,630 observations and 9 attributes, the message showed up:

"Excel cannot complete this task with available resource. Choose less data or close other applications."

4.1 Data Cleaning

Data cleaning has been performed in MS Excel and R. Mostly manually with the help of MS Excel tools such as Find/Replace, Missing Values Filter, and R to find out encoding issues (writing functions that report issues).

It resulted in creation of a file named `space_mission.xlsx` that I used for further data analysis.

4.1.1 Field: Location

- Algeria: Corrected the country for launch for - Brigitte, Hammaguir, Algeria, France. Removed France because the location is in Algeria, not in France (4 occurrences).
- Kiribati: Kiritimati Launch Area has Pacific Ocean, although it belongs to the Republic of Kiribati (36 occurrences)
- Barents Sea: it was launched in the Barents Sea Launch Area by Russian submarines (3 occurrences).
- Yellow Sea

4.1.2 Field: Date

- Added missing Date for "Rocket Lab Rocket Lab LC-1A, Mahia Peninsula, New Zealand 0:02:00 Electron/Curie
  A Data With Destiny"

4.1.3 Field: Time

- There are 127 missing values. I have tried to find first ten of them to no avail.

4.1.4 Field: MissionStatus

I added a numerical field that I used in csv later on:

Original MissionStatus	Numerical MissionStatus
Success	1
Failure	2
Partial Failure	3
Prelaunch Failure	4

4.1.5 Checks

Total number of observations: 4630

Field	Missing	Range/Values
Company	0	len: 2-16
Location	-	-
Site	23	len: 0-23
Facility	3	len: 0-49
State	2839	len: 0-16
Area	0	len: 3-13
Date	0	10/4/1957 - 7/29/2022
Time	127	0:00-23:59
Rocket	0	len: 3-24
Mission	0	len: 3-51
RocketStatus	0	Retired, Active
Price (millions of $)	3365	2.5 - 5,000

Field	Missing	Range/Values
MissionStatus	0	Success, Failure, Partial Failure, Prelaunch Failure

4.1.6 Result

I cleaned manually names from all characters that might be problematic for the databases or further processing.

I created the `space-mission.xlsx` file and exported `missions.csv` and `places.csv` that can be used in further processing.

- `places.csv` - contains part of the original location, i.e., facility, state, area
- `missions.csv` - contains RocketStatus and MissionStatus changed to numbers, it also has an additional field called ID of the mission.
- `missions-all-fields.csv` - contains numeric values of RocketStatus and MissionStatus, the location has four elements

References

[1]   M. Analytics, "Space Missions." Maven Analytics, Nov. 2022. Available: https://app.mavenanalytics.io/datasets