

Проект по дисциплине “Программирование на языке Python”

Анализ данных о полетах в космос

Выполнила:
Богателина Вера Алексеевна,
группа БСЦ 234

1. Обоснование выбора показателей

Цель проекта: при помощи языка python проанализировать данные о полетах в космос, построить рейтинги стран по разным показателям, обнаружить космонавтов-рекордсменов, узнать среднюю продолжительность полетов и посчитать другие описательные статистики, продемонстрировав владение базовыми навыками в программировании на python.

Тема проекта выбрана не случайно: космос всегда вызывал мой особый интерес, поэтому я и приняла решение в рамках проекта изучить статистику о космонавтах и полетах на орбиту.

Ремарка: готовых файлов с нужной мне статистикой я не нашла, поэтому информация была взята с [CBS Space News](#). Таблицы с необходимыми данными я формировала самостоятельно (единственный недочет CBS ресурса: данные были актуальны на 2018 год, однако это не критично в рамках моего проекта).

На сайте [CBS Space News](#) представлены данные о многих показателях: общий список всех космонавтов, информация о завершенных и текущих (на момент 2018 года) миссиях и экспедициях, орбитальные статистики, данные о выходе в открытый космос и тд. В рамках проекта я использовала далеко не все из представленных материалов. Первый файл содержит информацию о рекордсменах по продолжительности полетов, второй - о всех космонавтах, летавших на орбиту с 1961 года, третий - о провалах миссий и жертвах космоса. На мой взгляд, все три статистики в совокупности могут многое рассказать и дать примерное представление о состоянии космической сферы на 2018 год.

2. Нахождение статистик и интерпретация результатов

Первый файл:

В первом файле представлены данные о рекордах космонавтов по кумулятивной продолжительности полетов. На каждой строке таблицы содержится информация о позиции в рейтинге, имени и фамилии астронавта, далее в одной ячейке записывается через знак слеш ('/') количество дней, проведенных в космосе, и число полетов. Для начала, все данные были считаны в словарь, был заполнен список для продолжительности полетов (единица измерения преобразована из суток в часы). После этого, я посчитала медиану и среднюю длительность нахождения в космосе среди рекордсменов (поскольку файл изначально содержал статистику по рекордам). Затем были вычислены максимальная и минимальная продолжительность полета, и

благодаря максимальному значению был найден лидер среди космонавтов по суммарному количеству часов полета. На экран выводится его имя и фамилия, а также общее время пребывания в космосе (в часах) и число совершенных полетов.

Средняя продолжительность полетов = 11515.71 часа, медиана = 9408 часов, максимальное значение 19272 часов (за 6 полетов), минимальное = 8688 часов (за 3 полета). Среди рекордсменов, медианная продолжительность полета принадлежит Александру Волкову, а минимальная соответствует Джеффу Уильямсу.

В результате программа вычислила, что Сергей Крикалев установил рекорд, проведя в космосе суммарно 19272 часа (или 803 дня) в течение 6 полетов. Продолжительность полетов Крикалева превышает среднее значение времени у космонавтов на 7756.29 часа, а медиану на 9864 часа. Если слегка углубиться в вычисления, то можно увидеть, что в среднем в каждый из своих 6 полетов Сергей Крикалев проводил в космосе 3212 часов, то есть чуть меньше 134 дней за полет.

Второй файл:

Следующий файл представляет собой большой массив данных о всех космонавтах, начиная с момента первого полета человека в космос 12 апреля 1961 года. Данные из файла использовались для создания двух рейтингов стран: по количеству астронавтов и по числу полетов. Подробнее о рейтинга рассказано в пункте 3.

Итак, в первую очередь необходимо было считать данные в словарь. Строки в исходном файле содержали имя космонавта, количество совершенных им полетов в космос, в одной ячейке гендер и национальность, разделенные дефисом ('-'), и остальные столбцы содержали информацию о миссиях, в которых принимал участие астронавт. Для подсчета нужной мне статистики я использовала только первые 3 показателя (то есть данные о миссиях не были включены в исследование).

Далее были выполнены следующие технические операции: заполнение словарей для рейтинга стран по количеству астронавтов и полетов, для подсчета мужчин и женщин, составление списка с общим количеством полетов в космос и вычисление суммарного числа космонавтов (подробное описание хода заполнения всех данных и подсчетов см. в коде).

В итоге была получена информация о том, что всего с 1961 по 2018 год в космос летал 531 профессиональный космонавт, из них 475 человек - мужчины и 56 - женщины (в изначальных данных были представлены только те астронавты, кто уже завершил полет, то есть миссии, проводившиеся в настоящий момент (для 2018 года) не отражены в статистике). Любопытно, но несмотря на тот факт, что первая женщина-космонавт Валентина Терешкова совершила полет в космос еще в 1963 году (спустя всего 2 года после того, как возможность полетов в принципе была открыта), в 2018 году процентное соотношение женщин космонавтов составляет 10,55 % от общего числа астронавтов.

Суммарное количество полетов в космос составило 1196, значит, на каждого космонавта в среднем приходится по 2.25 полета.

В результате для каждой страны были подсчитаны количество космонавтов-граждан и число совершенных ими полетов. На основе этих данных были составлены два рейтинга стран, об интерпретации которых речь пойдет далее.

Третий файл:

Наконец, последний из выбранных файлов носит довольно печальный характер, ведь представленные в нем данные говорят о человеческих жертвах. Не все люди, перечисленные в таблице, скончались в космосе, данные свидетельствуют о всех трагически погибших космонавтах в результате испытаний, провалов миссий, ЧП, болезней и тд (например, в таблице указано имя Юрия Гагарина, жизнь которого прервалась в результате авиакатастрофы во время тренировочного полета).

Статистика содержит информацию о имени и фамилии астронавта, дате происшествия несчастного случая (=дата смерти космонавта) в формате мм/дд/гг, и краткое описание причины катастрофы. Помимо исходного файла, я составила список названий месяцев с января по декабрь. Данные были считаны в словарь, затем была преобразована дата: месяц из числового формата переведен в буквенный, а год из двузначного числа в четырехзначное. После этого был заполнен словарь с информацией о проваленных миссиях: ключи - дата и причина катастрофы, значения - количество пострадавших. Наконец, готовый словарь был отсортирован по значениям и записан в новый файл.

В итоге код выводит на экран информацию о первых 5 катастрофах, которые унесли наибольшее количество жизней. Сначала выводится фраза о провале миссии, затем дата, количество жертв и причина произошедшего.

К сожалению, первые места в этом траурном рейтинге занимают два происшествия, каждое из которых забрало на тот свет по 7 космонавтов. Первый несчастный случай: провал миссии STS-51L из-за отказа ракеты-носителя SRB, а второй: катастрофа на “Колумбии” по причине повторного разрыва и проникновения WLE. Третье место занимает ЧП на Союзе-11, который разгерметизировался во время возвращения на Землю и лишил жизни троих космонавтов.

Работа с третьим файлом, была довольно трудной. Хочется верить, что в будущем число погибших не будет стремительно расти, и космос не будет жесток к отважным людям.

3. Рейтинг стран

В общей сложности в рамках моего проекта были созданы 3 рейтинга, каждый из которых я записывала в отдельные файлы. Рейтинг о жертвах космоса я проинтерпретировала чуть выше, поэтому сейчас дам комментарии по первым двум.

Первый рейтинг посвящен выявлению стран-лидеров по количеству астронавтов, а второй ранжирует страны по суммарному количеству полетов в космос. Каждый рейтинг создавался из словарей, где ключи - страны, а значения - число астронавтов или полетов в космос (зависит от рейтинга). Словари отсортированы в порядке убывания, на экран выводятся первые 5 стран-лидеров (это сделано для большей наглядности и компактности вывода). Также в первом рейтинге выведено процентное количество астронавтов конкретной страны, а во втором рейтинге: процентное выражение числа полетов, совершенных космонавтами каждой страны. Процентные доли не записывались в файлы с рейтингами, они служат лишь для более показательной визуализации данных.

Любопытно, что топ-5 стран в обоих рейтингах почти одинаковые: первые три места занимают США, СССР и Россия, соответственно, далее четвертое и пятое место в рейтинге по числу космонавтов занимают Япония и Франция, а в рейтинге по суммарному количеству полетов - Франция и Канада. В целом, можно сделать вывод, что лидерами в космической индустрии по многим показателям стабильно являются США и СССР (нынешняя Россия). Это утверждение подкрепляется не только числовыми показателями, но и процентным выражением: 63% астронавтов - граждане США и почти 70% полетов в космос совершили именно американцы. Космонавты СССР и России суммарно составляют 21.5% от общего числа астронавтов, они также совершили почти 20% всех полетов человечества. Процентные доли других стран достаточно невелики, большая часть из них не превышает единицы.

Феномен столь активного участия России (в прошлом СССР) и США в освоении и развитии космической индустрии можно объяснить с исторической точки зрения: всем известно, что именно эти страны были первыми покорителями космоса, вставшими на такой опасный и захватывающий путь, открывший человечеству множество возможностей.

4. Итоги

Проделанная работа представляет собой довольно интересную статистическую базу, которая может быть крайне полезна при грамотном использовании.

Искренне выражаю надежду на то, что достижения наших соотечественников поднимут Россию на пьедестал покорителя космоса. Наша страна станет лидером, способным вести за собой человечество в светлое будущее. Пусть Юрий Алексеевич Гагарин гордится звездными потомками!