

# Кейс от ООО «Газпромнефть НТЦ» для студенческого чемпионата ГПН по data science



Со дня учреждения гильдии искателей приключений члены гильдии взяли бесчисленное множество поручений, для выполнения которых им приходилось сражаться с кровожадными монстрами, исследовать древние руины, искать сокровища, добывать редкие минералы, спасать нерадивых путешественников и делать многие другие вещи. Любой житель волшебного государства мог, предложив разумное вознаграждение, разместить поручение в Гильдии. Многие герои в ходе выполнения поручений попробовали себя в новых для себя ролях, использовали уникальные виды оружия. Гильдия всегда ответственно подходила к работе со своими клиентами, собирая отзывы о выполненных заказах: насколько заказчик удовлетворён сроками, в которые герой смог выполнить поручение, качеством работы героя и даже вежливостью героя.

Поток клиентов не иссякал, и со временем поручений стало так много, что герои перестали вдумчиво выбирать, за какое из них взяться, а просто брали первый попавшийся листок с заданием и шли выполнять его. Это привело к тому, что стали случаться первые провалы миссий: некоторые герои не могли справиться со слишком могучим монстром, другие не были достаточно опытными следопытами, чтобы найти пропавшего путешественника. Появились негативные отзывы от клиентов, некоторые из них даже отказывались платить!

Вскоре у членов гильдии появилась идея использовать волшебный древесный компьютер лесной академии – Аркашу для определения того, кто из героев наилучшим образом подойдёт для выполнения каждой из миссий. Кроме того, было решено, что для выполнения особенно сложных поручений можно привлекать не одного, а сразу нескольких героев. Например, кто-то из членов команды будет отвечать за поиск убежища монстра и отвлечение его внимания, а остальные будут заниматься нанесением урона и спасением похищенных монстром жителей. Для реализации этого амбициозного плана в Аркашу были загружены все когда-либо выданные поручения, а также послужной список каждого члена гильдии с отзывами от клиентов.



Помогите учёным лесной академии обучить Аркашу выдавать правильные рекомендации: научите его на основе истории выполнения поручений и владения героя теми или иными навыками формировать команды для выполнения поручений.

Поручения написаны обычными жителями, которые не слишком разбираются в монстрах и пещерах, поэтому описания похожих поручений могут очень сильно отличаться друг от друга: где-то каменный дракон назван ходящей горой, а кто-то назвал колодец, в котором видел кикимору, дыркой в земле. Возможно, для более качественного анализа Аркашу нужно научить объединять синонимичные формулировки.

В ходе выполнения поручений все члены гильдии в обязательном порядке ведут дневник, где записывают, какие действия им пришлось совершить для выполнения поручений, сколько времени это заняло и насколько сложно для них это было.



В команде, которую Аркаша может посоветовать для выполнения поручения, должно быть от одного до четверых героев. При этом задача Аркаши – обеспечить максимальную суммарную выручку гильдии. Если на каждое поручение отправлять по четыре героя, то героев на все поручения точно не хватит. С другой стороны, если команды будут слишком долго выполнять поручения, или вообще не справятся с ними, то приток денег в гильдию также сократится.

Используемые данные:

1. Список всех когда-либо выданных гильдии поручений – как выполненных, так и новых. В списке для каждого поручения приведено его описание, размер денежного вознаграждения за выполнение, затраченное на выполнение время.

Номер поручения	Заказчик	Дата поручения	Выполнено	Дата выполнения	Затрачено дней	Сумма вознаграж деня	Описание
11000	Анна	1051-08-11	да	1051-08-21	4	6 000	В лесу по дороге от пещеры звери нападают на людей. Нужно убить их. Осмотрите все возможные укрытия, чтобы найти зверей.
11001	Мария	1051-07-09	да	1051-09-02	2	20 000	В лесу по дороге от пещеры были замечены разбойники. Нужно убить их. Будьте осторожны, так как разбойники могут быть опасными.
11002	Эмилио	1053-10-05	да	1053-10-20	7	22 500	По дороге из деревни монстры похитили путников. Нужно освободить их. Осмотрите все возможные укрытия, чтобы найти похищенных.
11003	Бьянка	1052-11-24	да	1052-12-02	5	5 500	Недалеко от города видели монстров. Нужно победить их. Будьте осторожны, так как монстры могут быть опасными.
11004	Бьянка	1052-02-23	да	1052-03-30	8	10 500	В деревне у меня пропала сумка с документами. Нужно найти её как можно скорее.
11005	Анна	1051-09-07	да	1051-09-24	4	9 000	В лесу недалеко от города заметили зверей. Нужно убить их.
11006	Роман	1053-08-27	да	1053-10-18	8	8 000	В пещере появилось огромное каменное чудовище. Нужно уничтожить его.
11007	Иван	1051-02-14	да	1051-03-26	10	21 500	В пещере появилось огромное каменное чудовище. Нужно уничтожить его.
11008	Олег	1053-06-19	да	1053-06-24	3	12 000	Недалеко от города у меня потерялась драгоценность. Нужно найти её как можно скорее.
11009	Татьяна	1052-10-19	да	1052-12-13	2	28 500	В пещере появилось огромное каменное чудовище. Нужно убить его.
11010	Анна	1051-02-12	да	1051-02-18	4	22 500	Недалеко от города встречаются звери. Нужно убить их. Будьте осторожны, так как звери могут быть опасными.
11011	Аделина	1052-02-19	да	1052-02-28	3	6 500	Недалеко от города заметили разбойников. Нужно прогнать их. Будьте осторожны, так как разбойники могут быть опасными.
11012	Эмилио	1052-01-08	да	1052-02-14	4	11 000	В пещере завёлся дракон. Нужно его убить.
11013	Бабушка Синь	1053-05-23	да	1053-06-06	10	24 500	В лесу около деревни разбойники нападают на людей. Нужно проучить их.
11014	Татьяна	1052-05-16	да	1052-05-30	3	17 500	В лесу по дороге от пещеры заметили зверей. Нужно убить их.
11015	Анна	1052-12-25	да	1053-01-30	7	24 500	В лесу по дороге от пещеры заметили зверей. Нужно прогнать их. Осмотрите все возможные укрытия, чтобы найти зверей.
11016	Надя	1051-10-27	да	1051-12-02	4	26 500	В лесу около деревни видели монстров. Нужно уничтожить их. Осмотрите все возможные укрытия, чтобы найти монстров.
11017	Максимилиан	1053-03-03	да	1053-03-08	4	11 500	В пещере завёлся дракон. Нужно его уничтожить.
11018	Надя	1053-03-14	да	1053-05-13	9	19 500	По дороге из деревни монстры похитили путников. Нужно освободить их. Осмотрите все возможные укрытия, чтобы найти похищенных.
11019	Иван	1053-05-22	да	1053-06-08	10	7 500	В городе у меня пропала сумка с документами. Нужно найти её как можно скорее.
11020	Бьянка	1051-06-24	да	1051-07-14	6	15 500	Недалеко от города монстры похитили путников. Нужно спасти их. Осмотрите все возможные укрытия, чтобы найти похищенных.
11021	Чарли	1051-03-07	да	1051-04-04	5	13 000	В лесу по дороге от пещеры заметили монстров. Нужно победить их.
11022	Надя	1051-08-23	да	1051-08-29	2	26 000	В пещере появилось огромное каменное чудовище. Нужно убить его.

2. Дневники всех героев. В каждой записи указано, к какому из поручений она относится, а также в какой роли герой выполнял соответствующее действие, и сколько времени заняло выполнение.

Номер поручения	Герой	Запись в дневнике	Затрачено часов	Роль
11000	Мартин	разжечь костёр	1,0	рейнджер
11000	Мартин	выследить цель	6,0	следопыт
11001	Альфред	разжечь костёр	1,0	рейнджер
11001	Альфред	убить цель	3,0	мечник
11001	Альфред	залечить раны	18,0	лекарь
11002	Мартин	выследить цель	6,0	следопыт
11003	Бендер	убить цель	1,0	мечник
11004	Юлия	найти пропажу	4,5	следопыт
11004	Юлия	отыскать заказчика	3,0	следопыт
11005	Агата	разжечь костёр	2,0	рейнджер

3. Оценки, выставленные заказчиками каждому герою за каждое выполненное поручение по шкале от 2 до 5

Номер поручения	Герой	Оценка за качество	Оценка по срокам	Оценка за вежливость
11000	Мартин	4	3	4
11001	Альфред	5	5	4
11002	Мартин	5	4	4
11003	Бендер	2	4	3
11004	Юлия	4	4	5
11005	Агата	4	2	5
11006	Фредерик	4	4	4
11007	Альфред	5	3	2
11008	Фредерик	3	4	3
11009	Пастушок	2	4	5

### Формат решения:

Ссылка на GitHub, содержащая:

1. Файл Jupyter notebook с реализацией решения
2. Текстовое описание процесса выполнения задания с обязательным указанием принятых допущений (можно отдельным файлом, но лучше – как часть Jupyter notebook, чтобы текстовые описания шли перед каждым крупным блоком кода)
3. Таблица с командами героев, рекомендуемых для каждого из не выполненных поручений (отдельным файлом)
4. Любые другие таблицы, которые участник считает нужным предоставить (например, профили навыков героев)
5. Сравнение рекомендуемых Аркашей команд с несколькими вариантами случайно распределённых команд по полученному суммарному денежному вознаграждению и затраченному времени (отдельным файлом или как часть Jupyter notebook)

### Что мы будем искать, при проверке решений:

1. внимательность к данным
2. методичность и последовательность выводов
3. подходы к решению задачи и их реализация
4. аккуратность и читаемость кода
5. информативность отчёта