

Детектирование подозрительных личностей с помощью данных о совершенных полётах

Докладчик: Корст Ростислав

Команда: Корст Ростислав,
Лазичный Иван,
Кревский Михаил,
Сонин Андрей



План доклада

- Постановка задачи
- Используемые данные
- Методы решения
- Результаты
- Заключение



Постановка задачи



Пассажиры самолета



Постановка задачи



Пассажиры самолета



Подозрительная личность



Постановка задачи



Пассажиры самолета



Подозрительная личность

Наша задача – попробовать детектировать подозрительных личностей !



Используемые данные

- .csv
- .tab
- .xml
- .json
- .pdf
- .yaml
- zip



Используемые данные

- .csv
- .tab
- .xml
- .json
- .pdf
- .yaml
- zip



Используемые данные

- .csv →
- .tab →
- .xml →
- .json →
- .pdf
- .yaml
- zip

DATAFRAME1

DATAFRAME2

DATAFRAME3

DATAFRAME4



Используемые данные

- .csv →
- .tab →
- .xml →
- .json →
- .pdf
- .yaml
- zip

DATAFRAME1

DATAFRAME2

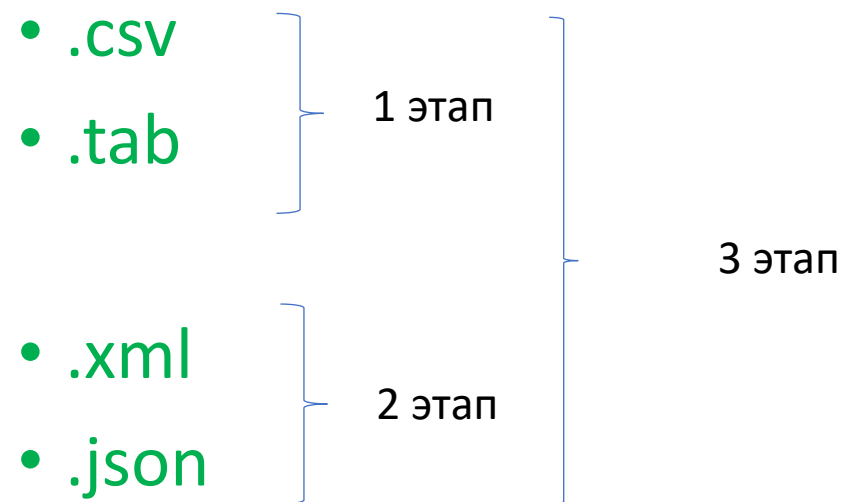
DATAFRAME3

DATAFRAME4

Как объединить эти данные ?
DATAFRAME_MAIN



Методы решения



Методы решения



1 ЭТАП ОБЪЕДИНЕНИЯ

1. Осуществили отображение русских ФИО в латинские для .tab
2. Объединили датафреймы .csv .tab по полям, соответствующим латинским ФИО и номеру рейса.
3. Использовали inner join для формирования результирующего датафрейма



Методы решения

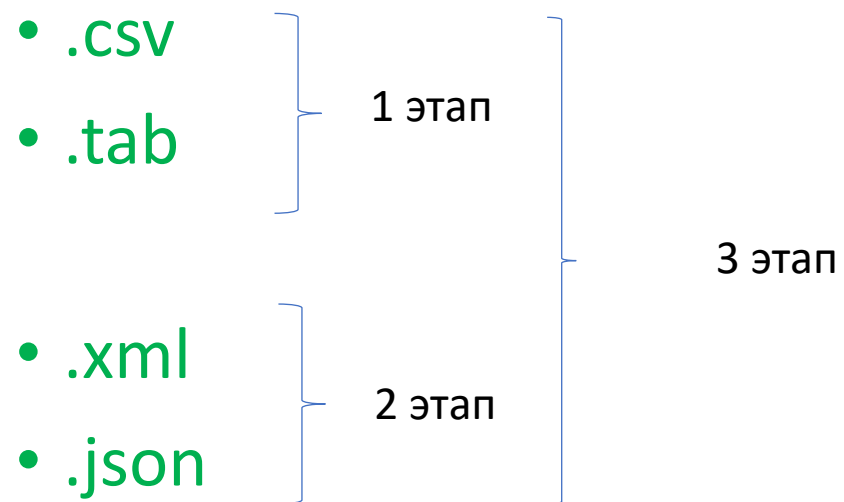


2 ЭТАП ОБЪЕДИНЕНИЯ

1. Объединили датафреймы .xml .json по полям, соответствующим карте лояльности, номеру рейса и дате полета.
2. Использовали inner join для формирования результирующего датафрейма



Методы решения



3 ЭТАП ОБЪЕДИНЕНИЯ

1. Объединили датафреймы по полям, соответствующим латинским ФИ, номеру рейса и дате полета.
2. Использовали inner join для формирования результирующего датафрейма

(Один человек может летать много раз, поэтому важно получить запись для конкретного человека и его конкретном перелета)



Результаты

- Получен объединенный датафрейм, использующий предоставленные источники данных
- Сформулирована гипотеза о том, что среди людей часто летающих в одном направлении в близком промежутке времени могут оказаться подозрительные личности



Заключение

- Разработать методику проверки выдвинутой гипотезы



Спасибо за Внимание!

