**1) Задача 1.**  
  
Данные транзакций позводяют нам определить только явные показатели - количество затрат, баланс карт и другие количественные показатели. Для лучшего понимания пользователей необходимо расширенные данные о пользователях - пол, наличие авто, семьи, детей, возможные интересы (спорт,бизнес,красота,культура и пр).  
  
Мы выгрузили датасет с 50 тысячами строк в формате - id карты, id пользователя, сумма транзакции, id банка, дата, описание транзакции, возможная категория и гео-данные. Id были захешированы (изменены с сохранением правильного сочетания).  
  
**Формальное условие и данные**  
  
По имеющимся данным транзакций пользователей пользователей приложения PingFin разработать алгоритм определения пола и интересов пользователя.  
 **Обработка данных на входе**Имеются следующие файлы:  
  
PingFin.json — указаны поля: id карты, id пользователя, сумма транзакции, id банка, дата, описание транзакции, возможная категория и гео-данные.  
      
**Оценка**  
Правильность выполнения задания определяются по списку id пользователей, по данным которых у нас есть точная информация:  
- пол  
- наличие авто  
- наличие семьи  
- наличие детей  
- интересы (спорт,бизнес,красота,культура и пр)  
  
**2) Задача 2.**  
На основе данных транзакций мы кластеризуем пользователей в зависимости от поведения и демонстрируем те или иные предложения от партнеров, основываясь на общем поведении/аномалиях поведения пользователей в кластере.  
  
Для анализа используется датасет с 50 тысячами строк в формате - id карты, id пользователя, сумма транзакции, id банка, дата, описание транзакции, возможная категория и гео-данные. Id были захешированы (изменены с сохранением правильного сочетания).  
  
**Формальное условие и данные**  
По имеющимся данным транзакций пользователей пользователей приложения PingFin разработать алгоритм прогнозирования следующих трат/поступлений  
для последующих рекомендаций.  
  
**Обработка данных на входе**  
Имеются следующие файлы:  
  
PingFin.json — указаны поля: id карты, id пользователя, сумма транзакции, id банка, дата, описание транзакции, возможная категория и гео-данные.  
      
**Оценка**  
Правильность выполнения задания определяются по списку id пользователей, по которым у нас есть данные по транзакциям после последней даты в датасете.