

Анализ меню ресторанов быстрого питания:



VS



McDonald's (далее McD)

Starbucks (далее SB)

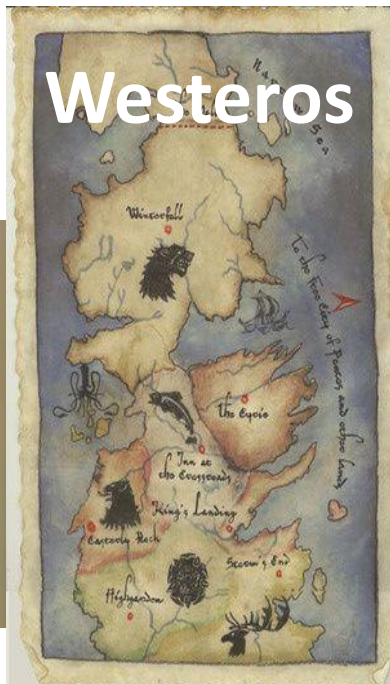
что выбрать для детского питания?

Westeros

Aleksandr Berezkin, Merr, Fatima Jandossova

Отчет подготовлен командой аналитиков «Westeros»
для детского развивающего центра «Мегамозг»

©Westeros, 2019

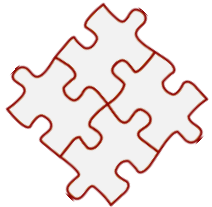


Задача

Выбрать какой ресторан быстрого питания лучше обеспечивает меню, удовлетворяющее основным требованиям:



Оптимальный баланс энергии и основных питательных веществ для детского питания



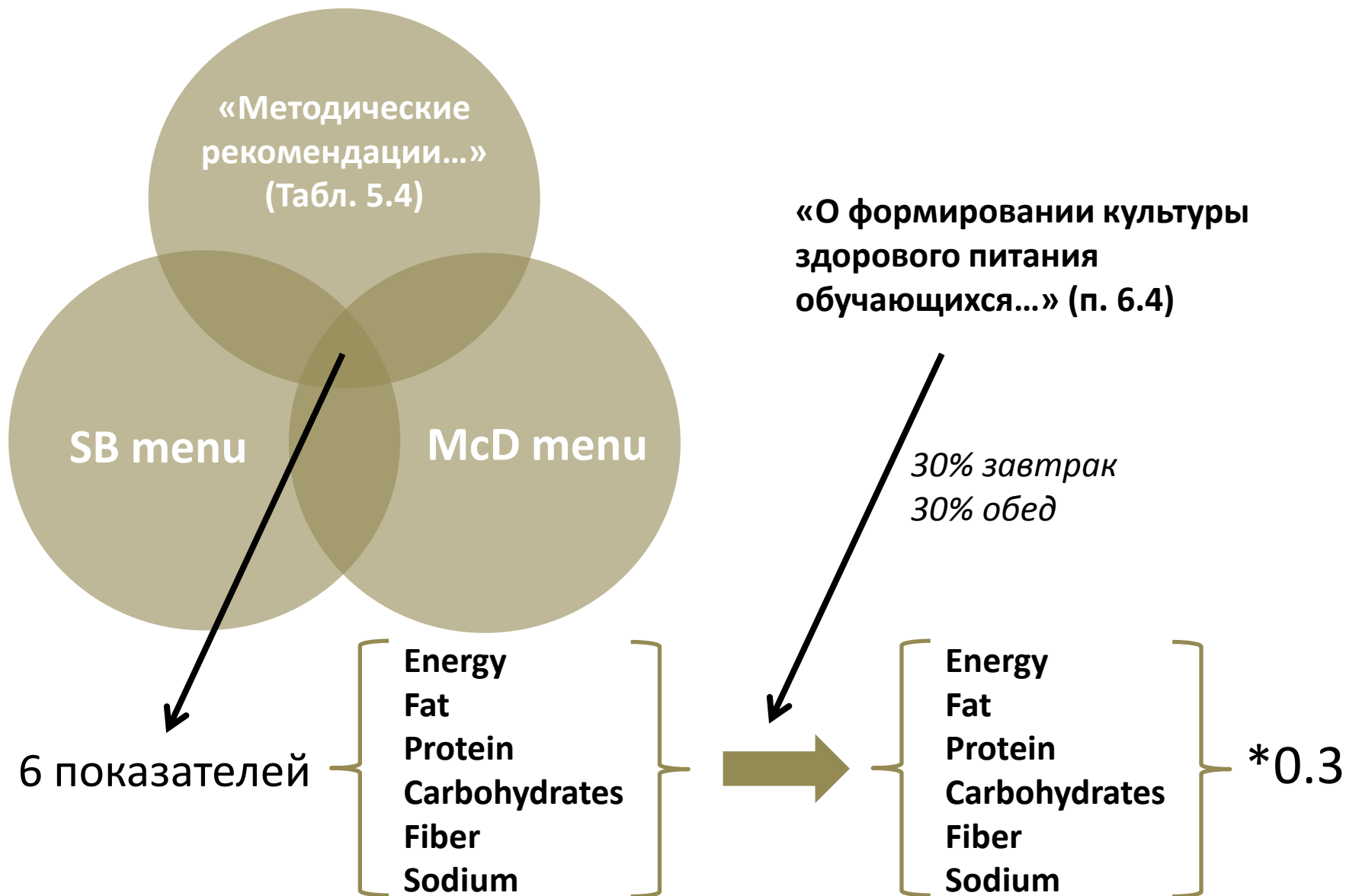
Возможность комбинировать разные блюда для обеспечения разнообразия питания



Контекст задачи и предоставленные данные:

- Летний детский лагерь, спортивные мероприятия
- Охват завтраков и обедов
- Лагерь действует на территории РФ
- Данные по конкретным меню McD и SB

Создание словаря норм



Предобработка данных

1. Создание общего датасета, новый признак ресторана: анализ представленности блюд по ресторанам, обогащение данных
2. Создание признака «категория блюд», анализ категорий, выделение новых категорий
3. Очистка датасета: удаление блюд, содержащих нули по всем 6 показателям и содержащих хотя бы один показатель $> 60\%$ от соответствующей дневной нормы
4. Редактирование названий: удаление знаков вроде ©, ™ и т.д.

```
merge dfs  
add "SB/McD"  
add 60+ SB items
```



```
add "Kind"
```



```
del all-zero items  
del items > 2*norm
```



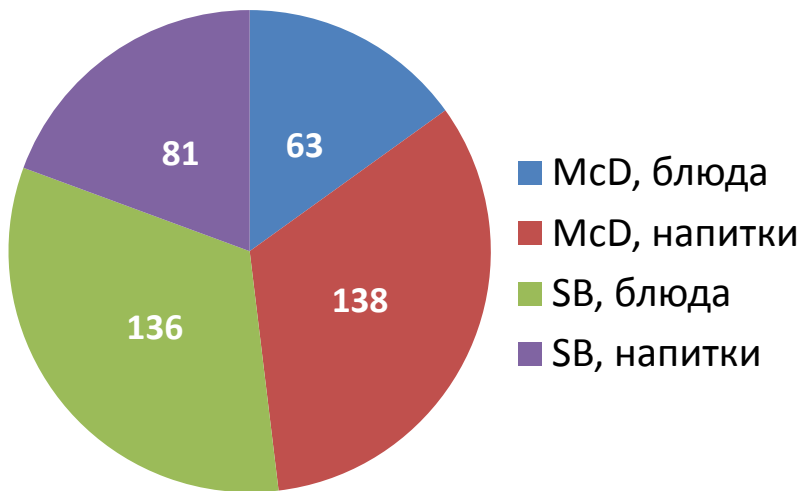
```
replace "©", "™" etc
```



Объединенный, сбалансированный датасет с 6 показателями по каждому блюду, не содержащий непригодных данных

Разведочный анализ

Доля блюд и напитков по ресторанам



- У SB больше еды, у McD - напитков

- **Калорийность и пищевая ценность:** максимальная у McD sweets/snacks. Эта же группа – лидер по **Углеводам**. Также высокая калорийность: SB Cold sandwiches (также лидируют по **Жирам**), SB Protein boxes (высокое содержание **Жиров**), McD Hot big breakfast

- **Распределение значений калорий и питательных веществ:** меню McD имеет в целом более сходные средние значения для разных категорий. В меню SB дисперсия выше, наблюдаются выбросы по средним значениям

В среднем блюда McD имеют более высокую **Калорийность**, содержание **Жиров, Углеводов, Белков и Натрия** ($McD:SB \approx 60:40$)

Среднее содержание **Пищевых волокон** в блюдах обоих ресторанов примерно одинаковое

Алгоритм анализа

Идея: эволюционный алгоритм – создание «осмысленных комбинаций» из меню каждого ресторана, выбор по критериям

Преимущества:

- Детализация: ответ на вопрос «Какой ресторан лучше» + конкретные варианты меню
- Гибкость: критерии можно менять под задачу, быстро получать результаты, сохранять и сравнивать их.
- Развитие и масштабируемость: данный подход может рассматриваться как задел, *proof of concept* для дальнейшего развития решения



Комбинации блюд:

- breakfast/dinner: **1-е, 2-е, напиток**
- light breakfast/dinner: **1-е, напиток**
- snack: **перекус, напиток**

Комбинации блюд (combination formats)

**breakfast
/dinner**



1-е,

2-е,

напиток



**light
breakfast
/dinner**



1-е,

напиток



snack

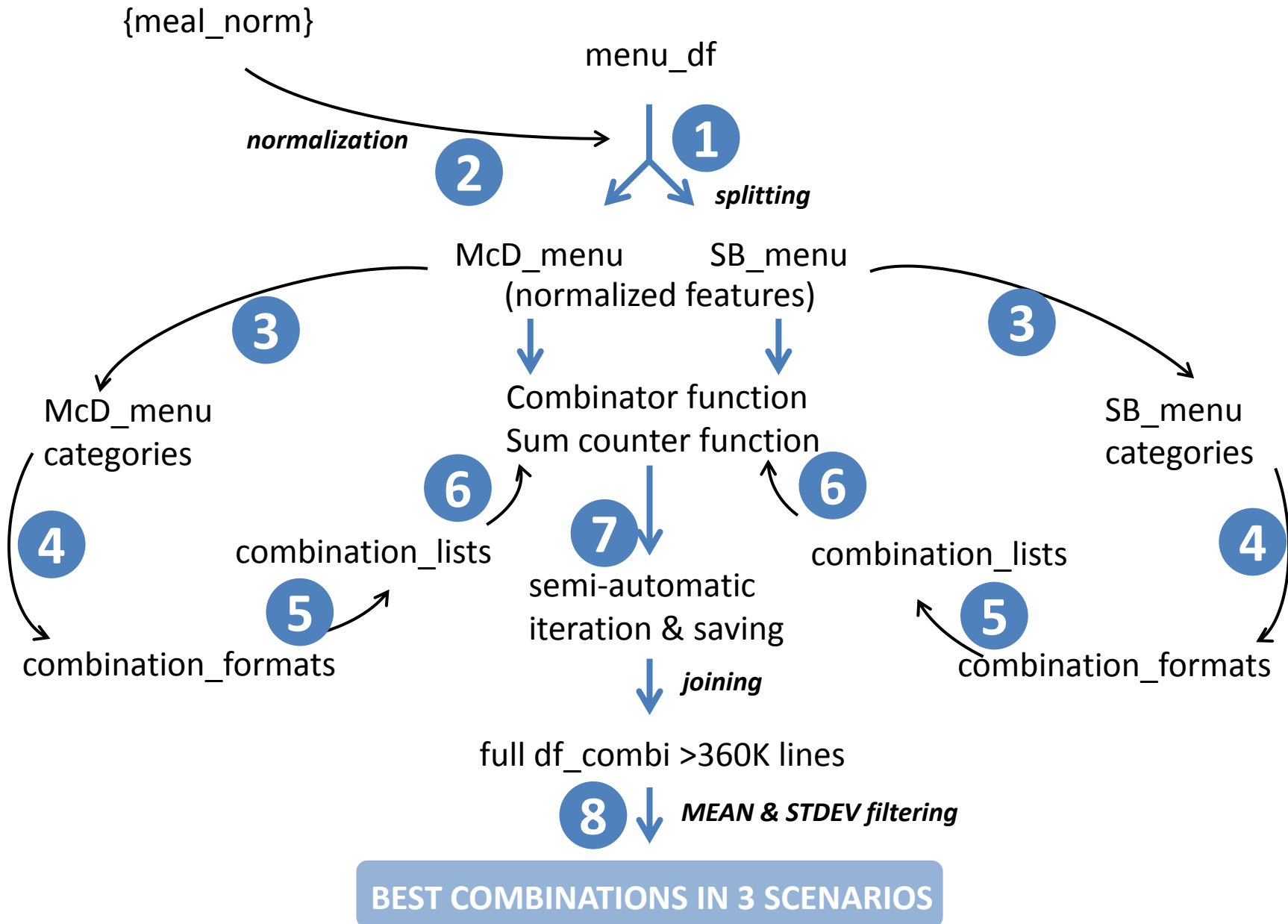


перекус,

напиток



Формальное представление алгоритма



Критерии и сценарии для фильтрации

MEAN: насколько среднее по 6 показателям близко к заданному числу

STDEV: насколько сильно могут отклоняться отдельные значения по показателям

	MEAN	STDEV	Energy
«Идеальный завтрак / обед»	1	0	-
«Жесткая норма»	1 ± 0.1	≤ 0.1	-
«Мягкая норма»	1 ± 0.2	≤ 0.2	-
«Спортивный день»	1.2 ± 0.1	≤ 0.1	≥ 1.2



Жесткая норма

Medium French Fries



Hot Fudge Sundae



Fat Free Chocolate Milk



Мягкая норма

Tomato & Mozzarella



Justin's Almond Butter



Tazo with Lemon



Спортивный день

Medium French Fries



Hot Fudge Sundae



Caramel Mocha (Medium)



Everything Cheese Bagel



Strawberry & Almonds
Grain



Cinnamon Frappuccino



Результаты

«Спортивный день»

MEAN = 1.2 ± 0.1
STDEV ≤ 0.1
Energy ≥ 1.2

McD_dinner	595
McD_light_dinner	371
SB_breakfast	4
SB_dinner	32

«Жесткая норма»

MEAN = 1 ± 0.1
STDEV ≤ 0.1

McD_breakfast	1
McD_dinner	217
McD_light_breakfast	6
McD_light_dinner	62
McD_snack	12
SB_breakfast	27
SB_dinner	196

«Мягкая норма»

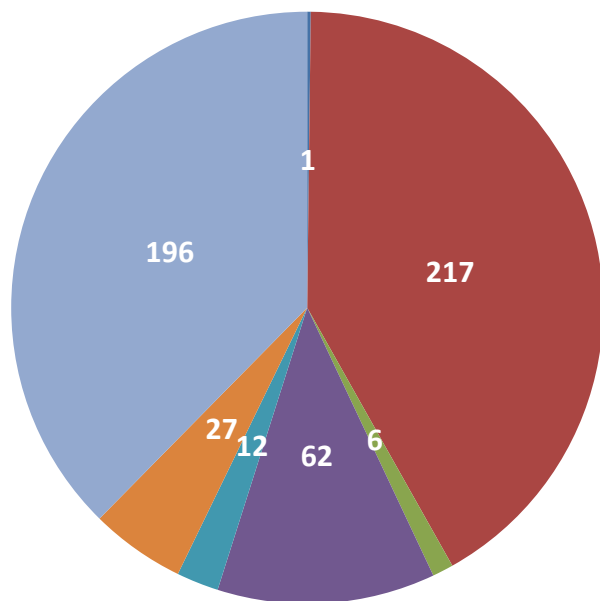
MEAN = 1 ± 0.2
STDEV ≤ 0.2

McD_breakfast	11
McD_dinner	1508
McD_light_breakfast	39
McD_light_dinner	973
McD_snack	71
SB_breakfast	515
SB_dinner	2410
SB_light_breakfast	2
SB_light_dinner	9

Из **360К** комбинаций отбор по «жесткой норме» прошли лишь **0,14%**, по «мягкой норме» - **1,43%**, а по «Спортивному дню» - **0,29%**



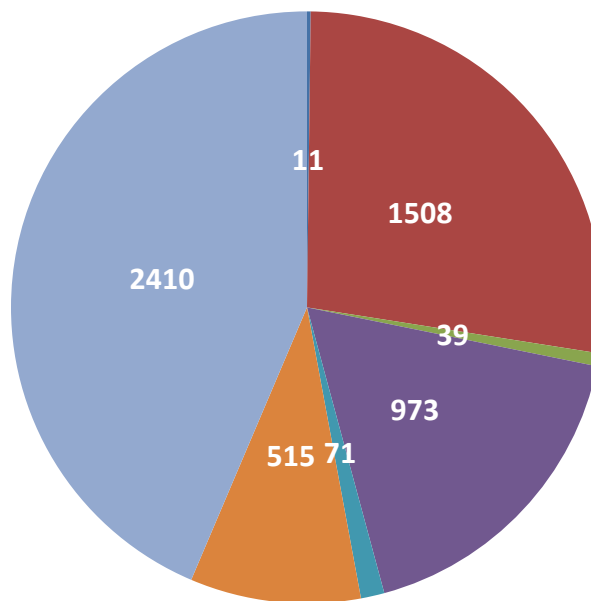
“Жесткая норма”



■ McD_breakfast
 ■ McD_dinner
 ■ McD_light_breakfast
 ■ McD_light_dinner
 ■ McD_snack
 ■ SB_breakfast
 ■ SB_dinner

Bcero SB: 223
 Bcero McD: 298

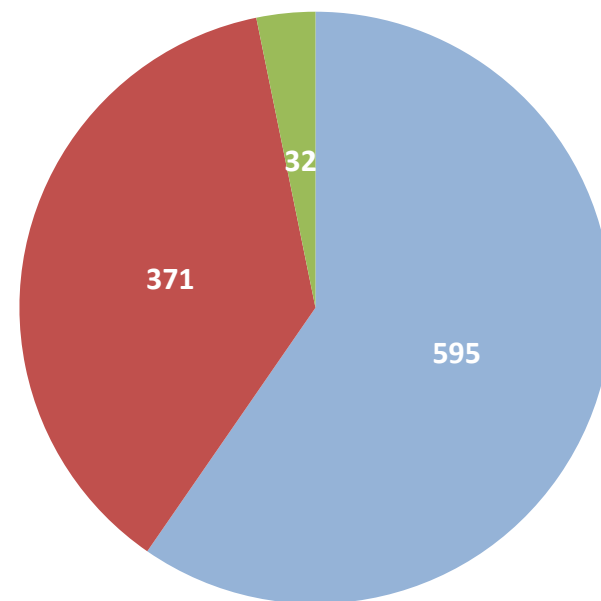
“Мягкая норма”



■ McD_breakfast
 ■ McD_dinner
 ■ McD_light_breakfast
 ■ McD_light_dinner
 ■ McD_snack
 ■ SB_breakfast
 ■ SB_dinner

Bcero SB: 2936
 Bcero McD: 2602

“Спортивный день”



■ McD_dinner
 ■ McD_light_dinner
 ■ SB_dinner

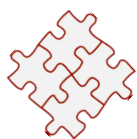
Bcero SB: 32
 Bcero McD: 996

Выводы: ответы на ключевые вопросы



Можно ли составить меню с оптимальным балансом энергии и основных питательных веществ для детского питания?

Да, возможно.



Можно ли составить меню, которое позволило бы комбинировать разные блюда для обеспечения разнообразия питания?

Возможно: в каждом сценарии множество комбинаций.



Какой из ресторанов больше подходит для цели клиента с учетом предоставленных им данных и в рассматриваемом контексте?



Приоритет:

- Питание, соответствующее нормам: **McD**
- Больше разнообразие блюд: **SB**
- Усиленный рацион для спортивных дней: **McD**

Ограничения анализа и рекомендации

1. Усредненная норма для девочек и мальчиков. Норма для мальчиков больше – мы взяли усредненную норму. Девочки переедают на 8%, мальчики недоедают на 8% (по энергии, жирам, углеводам и белкам).
2. Рекомендация клиента: оригинальные наборы данных McD и SB для США (прилагаются к ТЗ), тогда как лагерь «Мега Мозг» скорее всего, в России – мы взяли значения из нормативных документов РФ.
3. Наборы данных меню от клиента частично устарели. На сайтах ресторанов меню содержат много блюд, отсутствующих в данных клиента, при этом некоторые блюда из которых на сайтах также отсутствуют.



1. Реализация дополнительных сценариев, адаптированных под нормы для мальчиков и девочек
2. Рекомендуем клиенту уточнить место расположения детского лагеря и зафиксировать это в ТЗ
3. Алгоритм сбора данных с сайтов ресторанов на основе парсинга

Предложения по дальнейшей работе

1. Проверка других сценариев с различными фильтрами.
2. Включите данные о размерах порций и потенциальных аллергенах в продуктах
3. Поиск индексов питания или расчет интегрального индекса питания.
4. Улучшение кода. Code-review и полная автоматизация алгоритма. Обертывание всех действий в функции или классы.
5. Методы ML вместо ручных "сценариев": мы считаем, что методы оптимизации подойдут для данной задачи.
6. Генератор меню с различными настройками фильтров с GUI. Пользовательские датафреймы, список комбинаций, значения критериев для выбора лучших. Программа выбирает N случайных для составления уникальных меню.



!!!Полученные "лучшие комбинации" – лишь результат работы алгоритма и фильтрация данных по формальным критериям!!!

Приоритет горячих блюд. Warm sandwiches вместо Cold sandwiches для SB, Hot drinks вместо Cold drinks в обоих меню.
Учет уровня сахара. Особое внимание – мороженое McD.
Контроль уровня соли. Особое внимание уровню соли в продуктах, постоянный доступ к питьевой воде.

Использованные источники:

1. Меню ресторана [McDonalds](#)
2. Меню ресторана [Starbucks](#)
3. Табл. 5.4, «Методические рекомендации МР 2.3.1.2432-08:
[Нормы физиологических потребностей...»](#)
4. П. 6.4, Письмо Минобрнауки России от 12.04.2012 N 06-731
«О формировании культуры здорового питания обучающихся,
воспитанников» (вместе с [«Методическими рекомендациями...»](#))



Aleksandr Berezkin – поиск данных, рекомендации

Merr – алгоритм обработки, написание и документирование кода

Fatima Jandossova – поиск данных, подготовка презентации

Благодарим за внимание!
Ответим на вопросы