

# Анализ направлений туризма Тайланда



Назмиева Анастасия  
Юрьевна

# Цели и задачи проекта

## Цель

Проанализировать туристические направления в Тайланде и подобрать оптимальный вариант по отелю для отпуска в начале сентября 2025

## Задачи

Сравнить средние цены:

- по курортам (Пхукет vs Паттайя vs Као Лак)
- по датам (с 01.09-14.09)
- по рейтингу отелей
- по звёздам отелей

Выявить зависимости от метрик и определить топ-5 отелей для комфортного отдыха по соотношению цена-качество

# Этапы проекта

## Сбор данных

- Источником данных является сайт Anex Tour
- Парсинг данных с сайта информации по отелям с помощью библиотеки Selenium в Python

## Обработка данных

- Обработка данных с помощью библиотеки Pandas в Python
- Загрузка и обработка данных в БД SQLite
- Приведение атрибутов к одному naming convention
- Формирование информации по основным метрикам и параметрам (Source to target)

## Визуализация

- В качестве инструмента по визуализации выбран DataLens

## Анализ и выводы

- Анализ полученных данных
- Формирование выводов

# Парсинг данных







1 Импорт библиотек: Selenium – для сбора данных, Pandas – для обработки данных

2 Подключение к сайту <https://touristed.ru/toursearch>

```
#инициализация драйвера браузера.  
driver = webdriver.Chrome()  
#С помощью метода GET мы скажем браузеру, что надо открыть  
driver.get("https://touristed.ru/toursearch")  
time.sleep(10)  
departure_button = driver.find_element(By.CSS_SELECTOR, ".fancybox-button.fancybox-close-small")  
departure_button.click()  
time.sleep(5)
```

3 На сайте проставляем необходимые фильтры и раскрываем полный список.

В проекте используются следующие фильтры

<input checked="" type="radio"/> Туры с перелетом <input type="radio"/> Отели				
 ГОРОД ВЫЛЕТА Казань	 СТРАНА Таиланд	 ДАТЫ ВЫЛЕТА 1 сен - 14 сен	 НОЧЕЙ 5 - 10	 ТУРИСТЫ 2 взрослых
				

4 Парсим данные. Полный код представлен в [приложении 1](#)



# Обработка данных в Python и SQLite

Наименование поля  
(после парсинга данных)  
Пример данных

Процесс обработки данных

Наименование поля  
(после обработке данных)  
Пример данных

**name**  
Ibis Phuket Kata 3\*

Выделение **наименования отеля** в отдельный столбец

**hotel\_name**  
Ibis Phuket Kata 3\*

Выделение **звезды** отеля в отдельный столбец

**star**  
3

**region**  
Ката, Пхукет, 300 м до моря

Выделение **района** в отдельный столбец

**district**  
Ката

Выделение **города** отеля в отдельный столбец

**city**  
Пхукет

Выделение **кол-во метров до моря** в отдельный столбец

**meters\_sea**  
300 м до моря

**date\_tour**  
1 сен

Преобразование **даты тура** в формат ДД.ММ.ГГГГ

**date\_tour**  
01.09.2025

**full\_price**  
141 670 РУБ

Выделение **стоимости** в отдельный столбец.  
Преобразование в формат float

**price\_hotel**  
141 670

Выделение **валюты** в отдельный столбец

**currency**  
РУБ

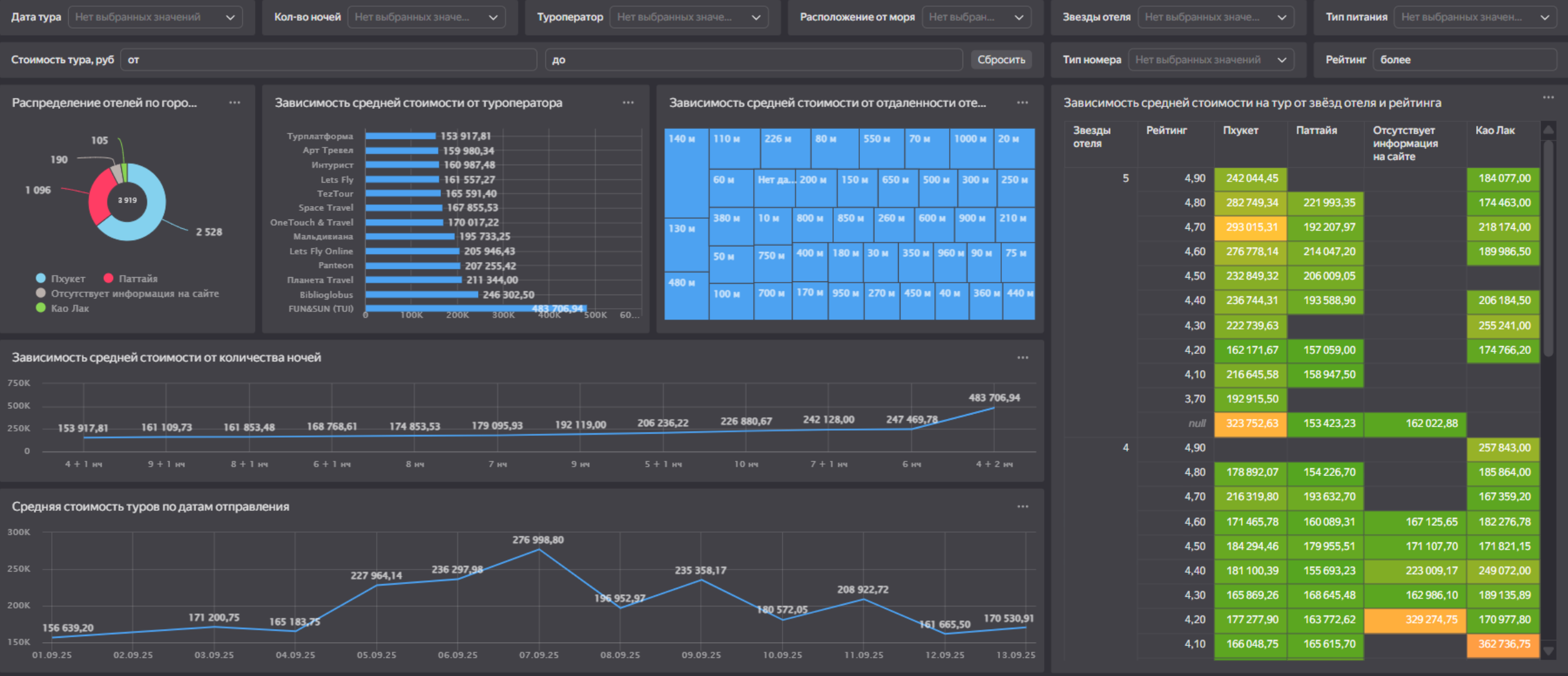
# Словарь данных

Таблица	Атрибут	Тип данных	Бизнес-определение	Пример данных
thailand	name	object	Наименование и количество звезд у отеля	Ibis Phuket Kata 3*
thailand	hotel_name	object	Наименование отеля	Ibis Phuket Kata
thailand	star	object	Количество звезд отеля	3
thailand	rating	object	Рейтинг отеля на сайте Anex tour	4.0
thailand	price	object	Минимальная стоимость тура по заданному отелю	141 670
thailand	description	object	Описание отеля	Уютный современный отель с удачным расположени...
thailand	region	object	Район, город и количество метров до моря у отеля	Ката, Пхукет, 300 м до моря
thailand	district	object	Район месторасположения отеля	Ката
thailand	city	object	Город месторасположения отеля	Пхукет
thailand	meters_sea	object	Количество метров до моря отеля	300 м до моря
thailand	href	object	Ссылка на отель с сайта Anex tour	<a href="https://tourcart.ru/hotel?cd=98899954#!/hotel=...">https://tourcart.ru/hotel?cd=98899954#!/hotel=...</a>
thailand	touroperator	object	Туроператор	Турплатформа
thailand	date_tour	datetime64[ns]	Дата вылета самолета	01.09.2025
thailand	nights	object	Количество ночей тура	4 + 1 нч
thailand	type_room	object	Тип комнаты тура	standard room
thailand	meal	object	Тип питания в отеле	ВВ - Только завтрак
thailand	full_price	object	Полная стоимость за тур	141 670 РУБ
thailand	price_hotel	float64	Стоимость тура	142 670
thailand	currency	object	Валюта тура	РУБ

# Визуализация (1/2)

[Возврат к слайду этапы проекта](#)

## Анализ направлений туризма Тайланда



С дашбордом можно ознакомиться по [ссылке](#)

# Визуализация (2/2)

Детализация

Наименование отеля с ссылкой на сайт	Рейтинг	Месторасположение отеля	Дата вылета	Кол-во ночей	Стоимость тура	Тип комнаты	Тип питания	Туроператор	Описание
<a href="#">The Pavilions Phuket 5*</a>	4,90	Банг Тао, Пхукет	05.09.2025	6 нч	268 933 РУБ	garden suite room	BB - Только завтрак	Biblioglobus	Отель в шаговой доступности от разнообразных достопримечательностей.
<a href="#">The Pavilions Phuket 5*</a>	4,90	Банг Тао, Пхукет	06.09.2025	6 нч	252 311 РУБ	tropical view suite	BB - Только завтрак	Biblioglobus	Отель в шаговой доступности от разнообразных достопримечательностей.
<a href="#">The Pavilions Phuket 5*</a>	4,90	Банг Тао, Пхукет	06.09.2025	6 нч	268 933 РУБ	garden suite room	BB - Только завтрак	Biblioglobus	Отель в шаговой доступности от разнообразных достопримечательностей.
<a href="#">The Pavilions Phuket 5*</a>	4,90	Банг Тао, Пхукет	07.09.2025	6 нч	253 675 РУБ	tropical view suite	BB - Только завтрак	Biblioglobus	Отель в шаговой доступности от разнообразных достопримечательностей.
<a href="#">The Pavilions Phuket 5*</a>	4,90	Банг Тао, Пхукет	08.09.2025	6 нч	252 311 РУБ	tropical view suite	BB - Только завтрак	Biblioglobus	Отель в шаговой доступности от разнообразных достопримечательностей.
<a href="#">The Pavilions Phuket 5*</a>	4,90	Банг Тао, Пхукет	08.09.2025	6 нч	268 933 РУБ	garden suite room	BB - Только завтрак	Biblioglobus	Отель в шаговой доступности от разнообразных достопримечательностей.
<a href="#">The Pavilions Phuket 5*</a>	4,90	Банг Тао, Пхукет	09.09.2025	6 нч	252 311 РУБ	tropical view suite	BB - Только завтрак	Biblioglobus	Отель в шаговой доступности от разнообразных достопримечательностей.
<a href="#">The Pavilions Phuket 5*</a>	4,90	Банг Тао, Пхукет	09.09.2025	6 нч	268 933 РУБ	garden suite room	BB - Только завтрак	Biblioglobus	Отель в шаговой доступности от разнообразных достопримечательностей.
<a href="#">Ramada Resort Khao Lak 4*</a>	4,90	Као Лак, 100 м до моря	01.09.2025	5 + 1 нч	257 843 РУБ	deluxe oasis	BB - Только завтрак	null	Отель прямо на побережье. Предлагает своим гостям отличный вид на море.
<a href="#">Amari Phuket (Ex. Amari Coral Beach Phuket) 4*</a>	4,80	Патонг, Пхукет, 50 м до моря	03.09.2025	7 нч	229 503 РУБ	superior balcony	BB - Только завтрак	Мальдивиана	Великолепный отель с большой зеленой территорией и прекрасным видом на море.
<a href="#">Amari Phuket (Ex. Amari Coral Beach Phuket) 4*</a>	4,80	Патонг, Пхукет, 50 м до моря	05.09.2025	6 нч	266 802 РУБ	superior balcony	BB - Только завтрак	Biblioglobus	Великолепный отель с большой зеленой территорией и прекрасным видом на море.
<a href="#">Amari Phuket (Ex. Amari Coral Beach Phuket) 4*</a>	4,80	Патонг, Пхукет, 50 м до моря	05.09.2025	6 нч	270 893 РУБ	superior ocean front balcony	BB - Только завтрак	Biblioglobus	Великолепный отель с большой зеленой территорией и прекрасным видом на море.
<a href="#">Amari Phuket (Ex. Amari Coral Beach Phuket) 4*</a>	4,80	Патонг, Пхукет, 50 м до моря	06.09.2025	6 нч	266 802 РУБ	superior balcony	BB - Только завтрак	Biblioglobus	Великолепный отель с большой зеленой территорией и прекрасным видом на море.



# Формирование выводов

## Распределение предложений по городам

**Пхукет** – **64,5% (2 528 шт.)** – развитая инфраструктура, международный аэропорт, разнообразие отелей (от бюджетных до премиум-класса), пляжи и активный туристический спрос.

**Паттайя** – **28% (1 096 шт.)** – близость к Бангкоку, доступные цены, насыщенная ночная жизнь, семейные и экскурсионные возможности.

**Као Лак** – **2,7% (105 шт.)** – более спокойный и эко-ориентированный курорт, подходит для дайвинга и отдыха вдали от шума.

По части предложений в отелях отсутствует информация на сайте 4,8% (190 шт)

**Пхукет** и **Паттайя** доминируют за счет массового туризма, тогда как **Као Лак** остается нишевым направлением.

## Основные факторы влияния стоимости на тур

Стоимость за тур с вылетом из Казани на даты с **1 сен – 14 сен** варьируется **от 119 до 489 тыс руб.** Данная зависимость устанавливается от следующими факторами:

### Местоположение отеля

- Наиболее бюджетные варианты в **городе с развитой инфраструктурой ночного досуга Паттайя**, но пляжи там менее привлекательны (мутная вода)
- Для пляжного отдыха лучше выбирать **Пхукет или Као Лак**.
- Чем дальше отель от моря, тем дешевле тур, но есть исключения. Например, отели в **непосредственной близости к пляжу** могут быть дешевле на 100 тыс руб., чем на удалении 130–200 м. до моря)

### Туроператор и Количество ночей

- **Турплатформы** предлагают самые выгодные цены на короткие туры (от **153 тыс руб. за 4+1 ночь**).
- Для болле продолжительного отдыха - **7 ночей** лидируют **Интурист** – 161 тыс руб и **Tez Tour** – 165 тыс руб.
- **FUN&SUN** выделяется высокой средней стоимостью (**483 тыс руб.**), что может быть связано с эксклюзивными предложениями и репутацией на рынке.

### Категория отеля (звездность) и рейтинг отеля

- Чем выше звездность и рейтинг, тем дороже тур.
- Незначительные отклонения присутствуют в средней стоимости между **3★ и 4★ отелями**.

### Даты вылета

- Наиболее выгодные даты – первая половина сентября (**1 сен - 4 сен**). Влияние низкого спроса после летних отпусков, в данное время года возможны дожди.

### Тип питания

- Туры **без питания** дешевле, но требуют дополнительных трат.
- Оптимальный вариант – **завтраки включены**.

# Формирование выводов

## ТОП-5 отелей по предпочтительным параметрам:

Наименование отеля	Рейтинг	Месторасположение отеля	Дата вылета	Кол-во ночей	Стоимость тура	Тип комнаты	Тип питания	Туроператор	Комментарий
<a href="#">Ravindra Beach Resort &amp; Spa 5*</a>	4,70	На Джомтьен, Паттайя, 50 м до моря	01.09.2025	4 + 1 нч	137 713 РУБ	superior	ВВ - Только завтрак	Турплатформа	<a href="#">Преимущества</a> Лучшее предложение для отдыха в Паттайи с хорошим рейтингом, в шаговой доступности море, завтрак включен в стоимость тура <a href="#">Недостатки</a> Кол-во ночей самое минимальное
<a href="#">My Beach Resort 5*</a>	4,90	Панва, Пхукет, 50 м до моря	01.09.2025	4 + 1 нч	165 587 РУБ	deluxe pool view	ВВ - Только завтрак	Турплатформа	<a href="#">Преимущества</a> На самую выгодную дату вылета -1 сен с отличным рейтингом, комната будет расположена с видом на бассейн и включен завтрак. <a href="#">Недостатки</a> Кол-во ночей самое минимальное
<a href="#">The Pavilions Phuket 5*</a>	4,90	Банг Тао, Пхукет	01.09.2025	4 + 1 нч	165 864 РУБ	tropical view suite	ВВ - Только завтрак	Турплатформа	<a href="#">Преимущества</a> Один из самых привлекательных отелей в Пхукете с развитой инфраструктурой и с хорошим рейтингом <a href="#">Недостатки</a> Кол-во ночей самое минимальное
<a href="#">Seaview Patong Hotel 4*</a>	4,30	Патонг, Пхукет, 100 м до моря	08.09.2025	9 нч	173 495 РУБ	superior	ВВ - Только завтрак	Интурист	<a href="#">Преимущества</a> Самый оптимальный вариант для продолжительного отдыха в Тайланде на 9 ночей <a href="#">Недостатки</a> Рейтинг и количество звезд у отеля ниже чем у остальных вариантов
<a href="#">Avani+ Khao Lak 5*</a>	4,90	Као Лак, 20 м до моря	03.09.2025	7 нч	184 077 РУБ	premier	ВВ - Только завтрак	Мальдивана	<a href="#">Преимущества</a> Один из бюджетных вариантов на среднее кол-во ночей - 7 с хорошим рейтингом и кол-вом звезд. Предназначен для спокойного отдыха <a href="#">Недостатки</a> отель находится в отдаленном месте – 73 км от аэропорта с не развитой городской инфраструктурой

# Приложение 1. Код парсинг данных (1/3)

```
# Создаем список для хранения данных
hotel_data = []

# Ожидаем загрузки карточек
info = WebDriverWait(driver, 30).until(
    EC.presence_of_all_elements_located((By.CLASS_NAME, "TVResultListItem")))
)
time.sleep(10)

for inf in info:
    try:
        # Прокручиваем к элементу
        driver.execute_script("arguments[0].scrollIntoView({block: 'center', behavior: 'smooth'});", inf)
        time.sleep(10)

        # Создаем словарь для хранения данных текущего отеля
        hotel = {
            'name': None,
            'rating': None,
            'price': None,
            'description': None,
            'region': None,
            'href': None,
            'tours': [] # Список для хранения всех вариантов туров
        }

        # Основные данные отеля
        try:
            hotel['name'] = inf.find_element(By.CLASS_NAME, "TVHotelInfoTitleLink").text
            hotel['rating'] = inf.find_element(By.CLASS_NAME, "TVResultItemBeforeDescription").text
            hotel['price'] = inf.find_element(By.CLASS_NAME, "TVResultItemPriceValue").text
            hotel['description'] = inf.find_element(
                By.CSS_SELECTOR, ".TVResultItemDescription.TVLineClampEnabled.TVLineClamp-M").text
            hotel['region'] = inf.find_element(By.CLASS_NAME, "TVResultItemSubTitle").text
            hotel['href'] = inf.find_element(By.CSS_SELECTOR, "div.TVResultItemTitle a[data-tvtourlink]").get_attribute("href")
        except NoSuchElementException as e:
            print(f"Ошибка в основных данных отеля: {str(e)}")
            continue
```

# Приложение 1. Код парсинг данных (1/3)

```
# Поиск всех вариантов туров
try:

    try:
        arrow = inf.find_element(By.CSS_SELECTOR, ".TVResultItemPrice .TVResultItemPriceValueArrow")
        driver.execute_script("arguments[0].click();", arrow)
        time.sleep(3) # Даем время для раскрытия списка
    except NoSuchElementException:
        print("Не найдена стрелка для раскрытия списка туров")

# Находим контейнер с турами
tours_container = inf.find_element(
    By.CSS_SELECTOR, ".TVHotelResultItem .TVHotelResultItemTours")

# Находим все строки с вариантами туров
tour_rows = tours_container.find_elements(
    By.CSS_SELECTOR, "t-tr.TVTourResultItem")

for row in tour_rows:
    try:
        tour_info = {
            'touroperator': row.find_element(
                By.CSS_SELECTOR, "div.TVTourResultItemOperator").get_attribute("title"),
            'date_tour': row.find_element(
                By.CSS_SELECTOR, ".TVTourResultItemDate").get_attribute("textContent"),
            'nights': row.find_element(
                By.CSS_SELECTOR, ".TVTourResultItemNights").get_attribute("textContent"),
            'type_room': row.find_element(
                By.CSS_SELECTOR, "t-span.TVTourResultItemRoom").get_attribute("textContent"),
            'meal': row.find_element(
                By.CSS_SELECTOR, ".TVTourResultItemMeal").get_attribute("textContent"),
            'full_price': row.find_element(
                By.CSS_SELECTOR, ".TVTourResultItemPriceValue").get_attribute("textContent") + " " +
                row.find_element(
                    By.CSS_SELECTOR, ".TVTourResultItemPriceCurrency").get_attribute("textContent")
        }
        hotel['tours'].append(tour_info)

    except NoSuchElementException as e:
        print(f"Не удалось извлечь некоторые данные тура: {str(e)}")
        continue
```



# Приложение 1. Код парсинг данных (1/3)

```
# Создаем расширенный DataFrame с отдельными строками для каждого варианта тура
```

```
expanded_data = []
```

```
for hotel in hotel_data:
```

```
    if not hotel['tours']:
```

```
        # Добавляем отель даже без туров
```

```
        expanded_data.append({
```

```
            'name': hotel['name'],
```

```
            'rating': hotel['rating'],
```

```
            'price': hotel['price'],
```

```
            'description': hotel['description'],
```

```
            'region': hotel['region'],
```

```
            'href': hotel['href'],
```

```
            'touroperator': 'Нет данных',
```

```
            'date_tour': 'Нет данных',
```

```
            'nights': 'Нет данных',
```

```
            'type_room': 'Нет данных',
```

```
            'meal': 'Нет данных',
```

```
            'full_price': 'Нет данных'
```

```
        })
```

```
    else:
```

```
        for tour in hotel['tours']:
```

```
            expanded_data.append({
```

```
                'name': hotel['name'],
```

```
                'rating': hotel['rating'],
```

```
                'price': hotel['price'],
```

```
                'description': hotel['description'],
```

```
                'region': hotel['region'],
```

```
                'href': hotel['href'],
```

```
                'touroperator': tour.get('touroperator', 'Нет данных'),
```

```
                'date_tour': tour.get('date_tour', 'Нет данных'),
```

```
                'nights': tour.get('nights', 'Нет данных'),
```

```
                'type_room': tour.get('type_room', 'Нет данных'),
```

```
                'meal': tour.get('meal', 'Нет данных'),
```

```
                'full_price': tour.get('full_price', 'Нет данных')
```

```
            })
```

```
# Создаем DataFrame
```

```
df = pd.DataFrame(expanded_data)
```

```
df
```