```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
%matplotlib inline
from sklearn import tree
```

```
data_orders = pd.read_csv('https://stepik.org/media/attachments/lesson/17834/orders.csv')
```

```
data_items = pd.read_csv('https://stepik.org/media/attachments/lesson/17834/items.csv')
```

```
data_sellers = pd.read_csv('https://stepik.org/media/attachments/lesson/17834/sellers.csv')
```

```
data_review = pd.read_csv('https://stepik.org/media/attachments/lesson/17834/order_reviews.csv')
```

	product_id	name_rus	CURRENCY
0	1	Центы Полосатый длинн	USD
1	2	Перчатки Клевый цветные	RUB
2	3	Перчатки Классные мужские	RUB
3	4	Перчатки Классные Женщ	RUB
4	5	Перчатки Классные ОСТЕ	RUB
5	6	Носки Полосатые муж. - 41	RUB
6	7	Носки Частные муж. - 53	RUB
7	8	Носки Модные Женщ. - 36	RUB
8	9	Носки Модные муж. - 249	RUB
9	10	Носки Классные Сатина	RUB
10	11	Гольфы длинные Сексуальн	RUB
11	12	Гольфы длинные Лапка Лис	RUB
12	13	Сумка для носков - 510	RUB
13	14	Чай, чайник, чайник-заварник	RUB

```
data_orders_order_id.nunique()
```

```
data_orders_order_id.count()
```

```
data_items.product_id.nunique()
```

```
data_items.product_id.count()
```

	order_id	customer_id	order_status	order_purchase_timestamp	order_approved_at	order_delivered_carrier_date	order_delivered_customer_date	order_estimated_delivery_date
0	2017-10-02 10:56:33	1. Получен	Получен	2017-10-02 10:56:33	2017-10-02 11:07:15	2017-10-04 19:55:00	2017-10-10 21:25:13	2017-10-18 00:00:00
1	2017-10-26 12:55:51	2. Доставлен	Доставлен	2017-10-26 12:55:51	2017-10-26 13:05:13	2017-10-27 15:48:00	2017-10-30 17:14:25	2017-11-06 00:00:00
2	2017-10-16 13:36:05	3. Доставлен	Доставлен	2017-10-16 13:36:05	2017-10-16 13:45:35	2017-10-17 11:14:00	2017-10-18 14:35:12	2017-10-31 00:00:00
3	2017-10-16 16:46:17	4. Доставлен	Доставлен	2017-10-16 16:46:17	2017-10-17 03:22:32	2017-10-18 19:44:00	2017-10-20 21:18:39	2017-11-29 00:00:00
4	2017-10-16 16:16:42	5. Доставлен	Доставлен	2017-10-16 16:16:42	2017-10-16 16:30:30	2017-10-18 19:30:00	2017-10-20 21:54:00	2017-11-29 00:00:00
5	2017-10-16 18:39:31	6. Доставлен	Доставлен	2017-10-16 18:39:31	2017-10-16 18:50:18	2017-10-18 14:05:00	2017-10-19 18:43:53	2017-10-31 00:00:00
6	2017-10-16 08:38:13	7. Отменен	Отменен	2017-10-16 08:38:13	NaN	NaN	NaN	2017-11-06 00:00:00
7	2017-10-16 08:38:13	8. Доставлен	Доставлен	2017-10-16 08:38:13	2017-10-16 08:50:18	2017-10-18 17:30:00	2017-10-19 00:28:06	2017-11-06 00:00:00
8	2017-10-16 08:38:13	9. Доставлен	Доставлен	2017-10-16 08:38:13	2017-10-16 08:50:18	2017-10-18 17:30:00	2017-10-19 00:28:06	2017-11-06 00:00:00
9	2017-10-16 08:38:13	10. Доставлен	Доставлен	2017-10-16 08:38:13	2017-10-16 08:50:18	2017-10-18 17:30:00	2017-10-19 00:28:06	2017-11-06 00:00:00
10	2017-10-16 08:38:13	11. Доставлен	Доставлен	2017-10-16 08:38:13	2017-10-16 08:50:18	2017-10-18 17:30:00	2017-10-19 00:28:06	2017-11-06 00:00:00
11	2017-10-16 08:38:13	12. Доставлен	Доставлен	2017-10-16 08:38:13	2017-10-16 08:50:18	2017-10-18 17:30:00	2017-10-19 00:28:06	2017-11-06 00:00:00
12	2017-10-16 08:38:13	13. Доставлен	Доставлен	2017-10-16 08:38:13	2017-10-16 08:50:18	2017-10-18 17:30:00	2017-10-19 00:28:06	2017-11-06 00:00:00
13	2017-10-16 08:38:13	14. Доставлен	Доставлен	2017-10-16 08:38:13	2017-10-16 08:50:18	2017-10-18 17:30:00	2017-10-19 00:28:06	2017-11-06 00:00:00

```
data_review.head(14)
```

```
data_orders_order_id.nunique()/data_orders_order_id.count()
```

```
data_items.product_id.nunique()/data_items.product_id.count()
```

```
data_orders_order_id.count()/data_items.product_id.count()
```

```
data_review.review_score.hist(bins=5, figsize=(10,5))
```

```
data_items.price.hist(bins=50, figsize=(10,5))
```

```
data_items.price.hist(bins=50, figsize=(10,5))
```

```
data_items.price.hist(bins=50
```