

Перезапуск Data Engineer

Решаем тест по предыдущей лекции
до 18.40

Работа с данными с помощью Python



Оснoвы Pandas

SQL	Pandas
<code>select * from airports</code>	<code>airports</code>
<code>select * from airports limit 3</code>	<code>airports.head(3)</code>
<code>select id from airports where ident = 'KLAX'</code>	<code>airports[airports.ident == 'KLAX'].id</code>
<code>select distinct type from airport</code>	<code>airports.type.unique()</code>

Оснoвы Pandas

SQL	Pandas
<pre>select * from airports where iso_region = 'US-CA' and type = 'seaplane_base'</pre>	<pre>airports[(airports.iso_region == 'US-CA') & (airports.type == 'seaplane_base')]</pre>
<pre>select ident, name, municipality from airports where iso_region = 'US-CA' and type = 'large_airport'</pre>	<pre>airports[(airports.iso_region == 'US-CA') & (airports.type == 'large_airport')][['ident', 'name', 'municipality']]</pre>

Оснoвы Pandas

SQL	Pandas
<pre>create table heroes (id integer, name text);</pre>	<pre>df1 = pd.DataFrame({'id': [1, 2], 'name': ['Harry Potter', 'Ron Weasley']})</pre>
<pre>insert into heroes values (1, 'Harry Potter');</pre>	<pre>df2 = pd.DataFrame({'id': [3], 'name': ['Hermione Granger']})</pre>
<pre>insert into heroes values (2, 'Ron Weasley');</pre>	
<pre>insert into heroes values (3, 'Hermione Granger');</pre>	<pre>pd.concat([df1, df2]).reset_index(drop=True)</pre>

Кодинг - это:

- 1% действительно кодинг
- 40% отладка
- 15% перерыв на кофе
- 30% поиск ошибок в google
- 9% смотрите с коллегами на экран
- 5% попытки копировать/вставлять решения из Stack Overflow