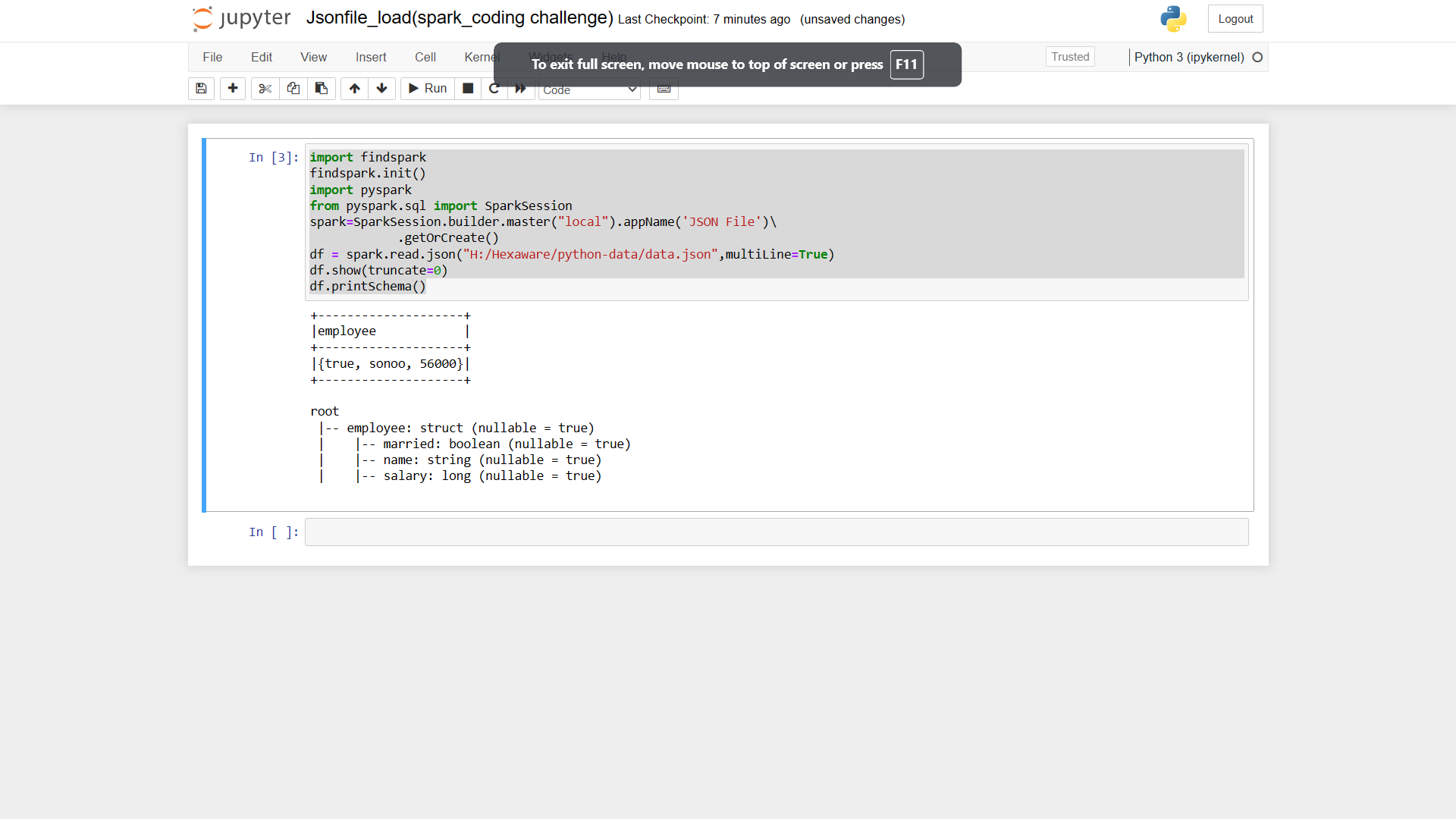
**Name : Hemalatha Karumuri**

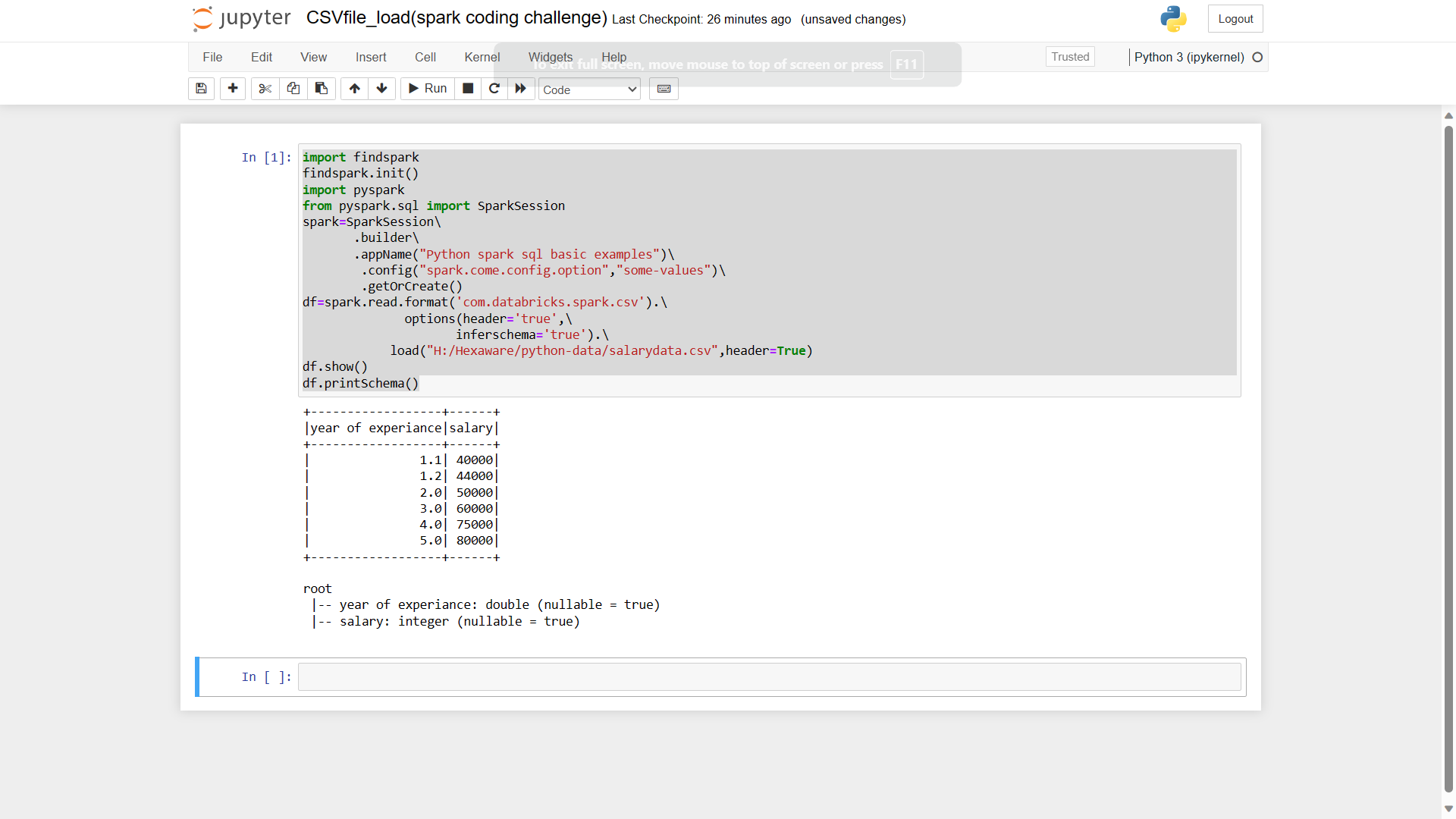
**Date : 27-12-23**

**SPARK CODING CHALLENGE**

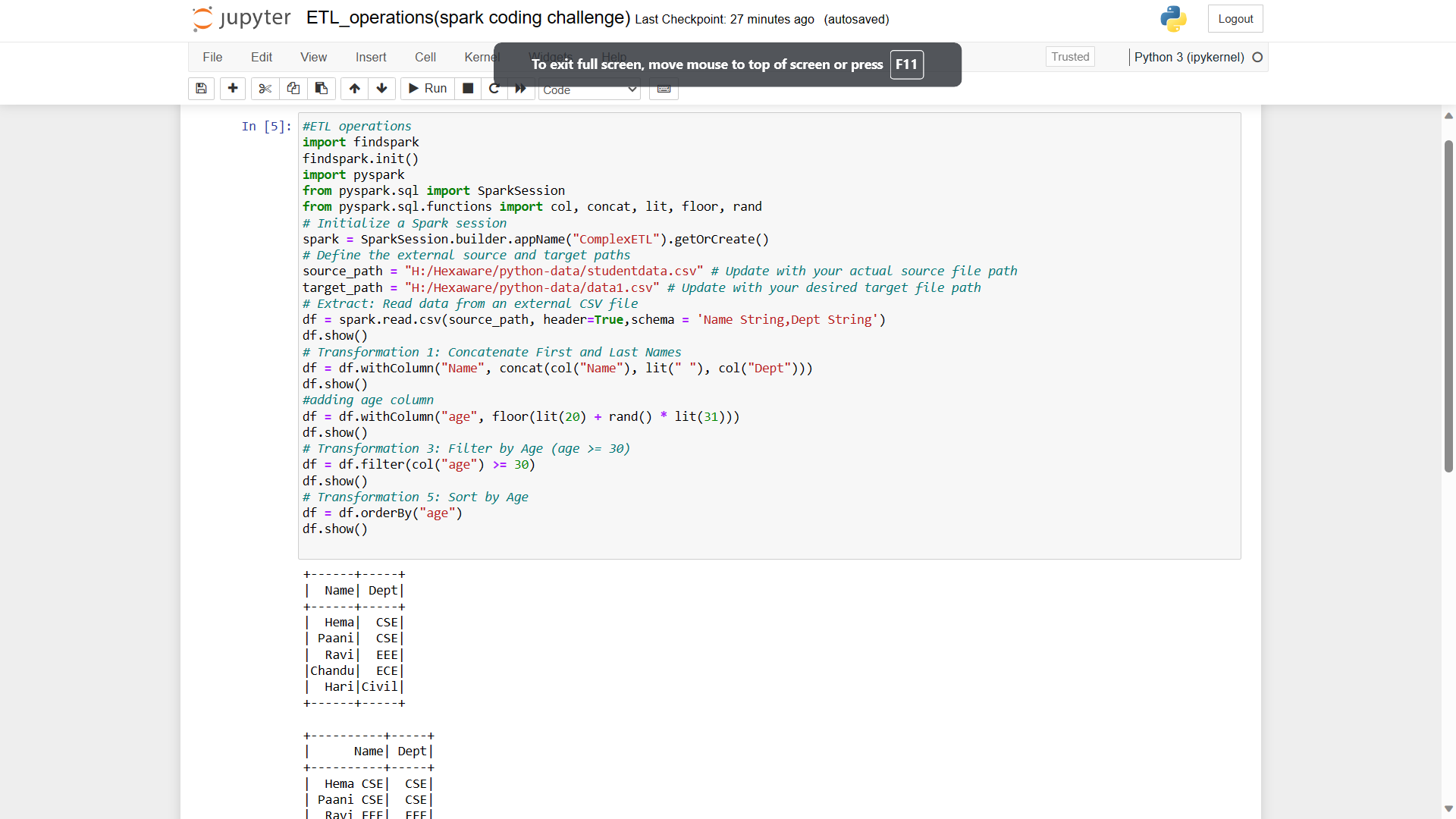
**1-A)Implement Processing JSON file**

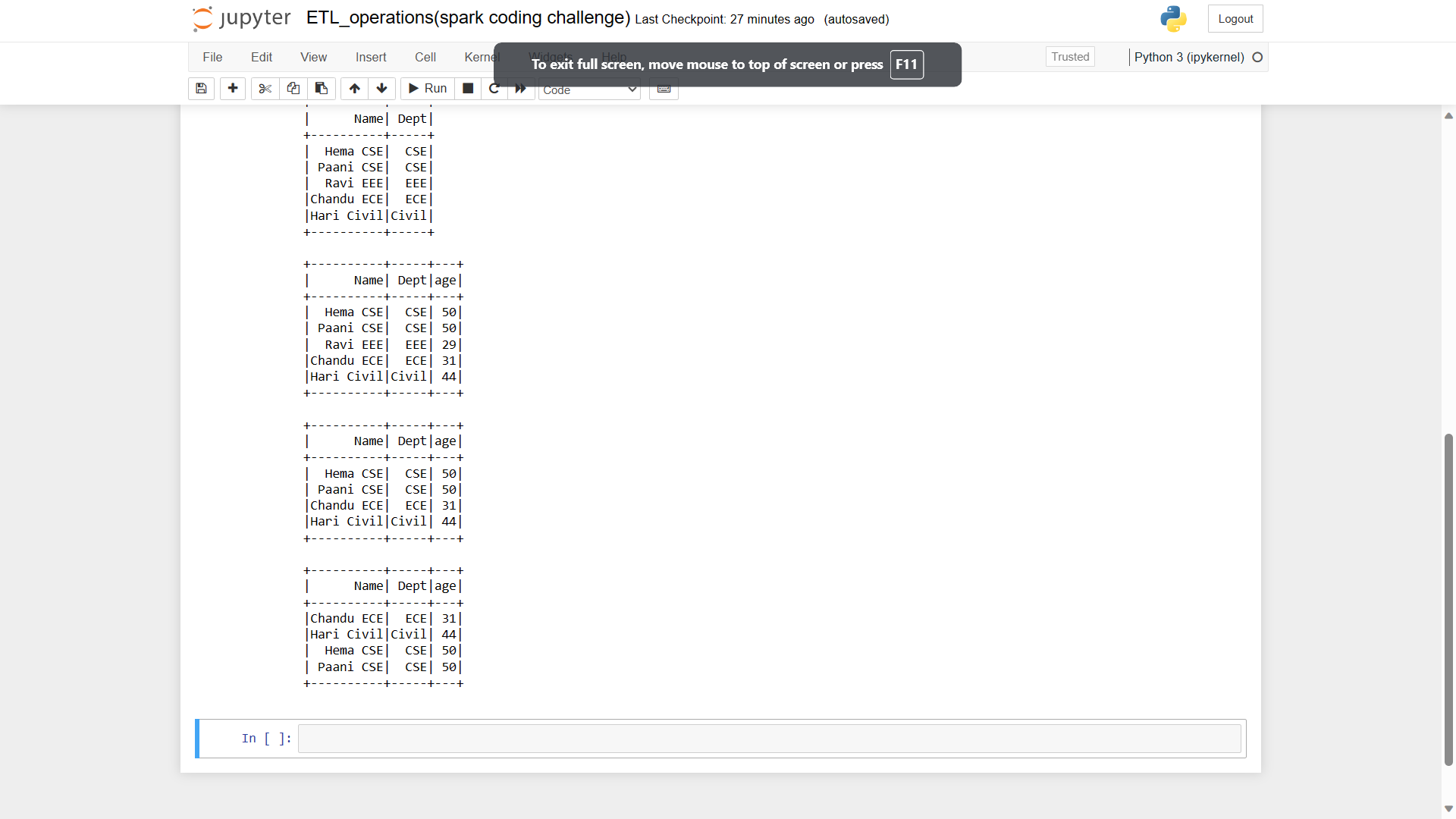


**1-b)Implement Processing CSV file data with Pyspark**

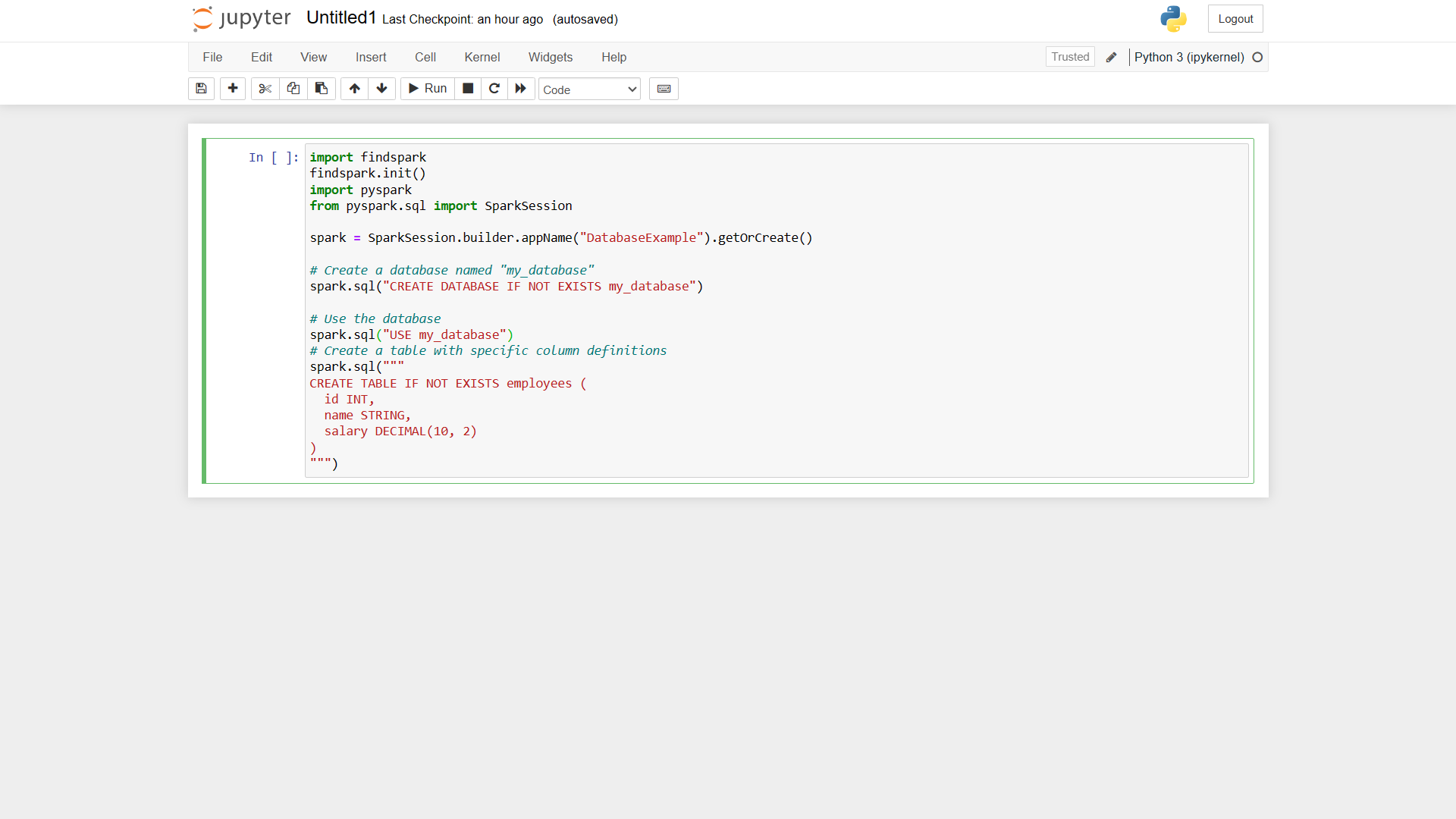


**2) Explain ETL (Extract, Transform, Load) with PySpark**



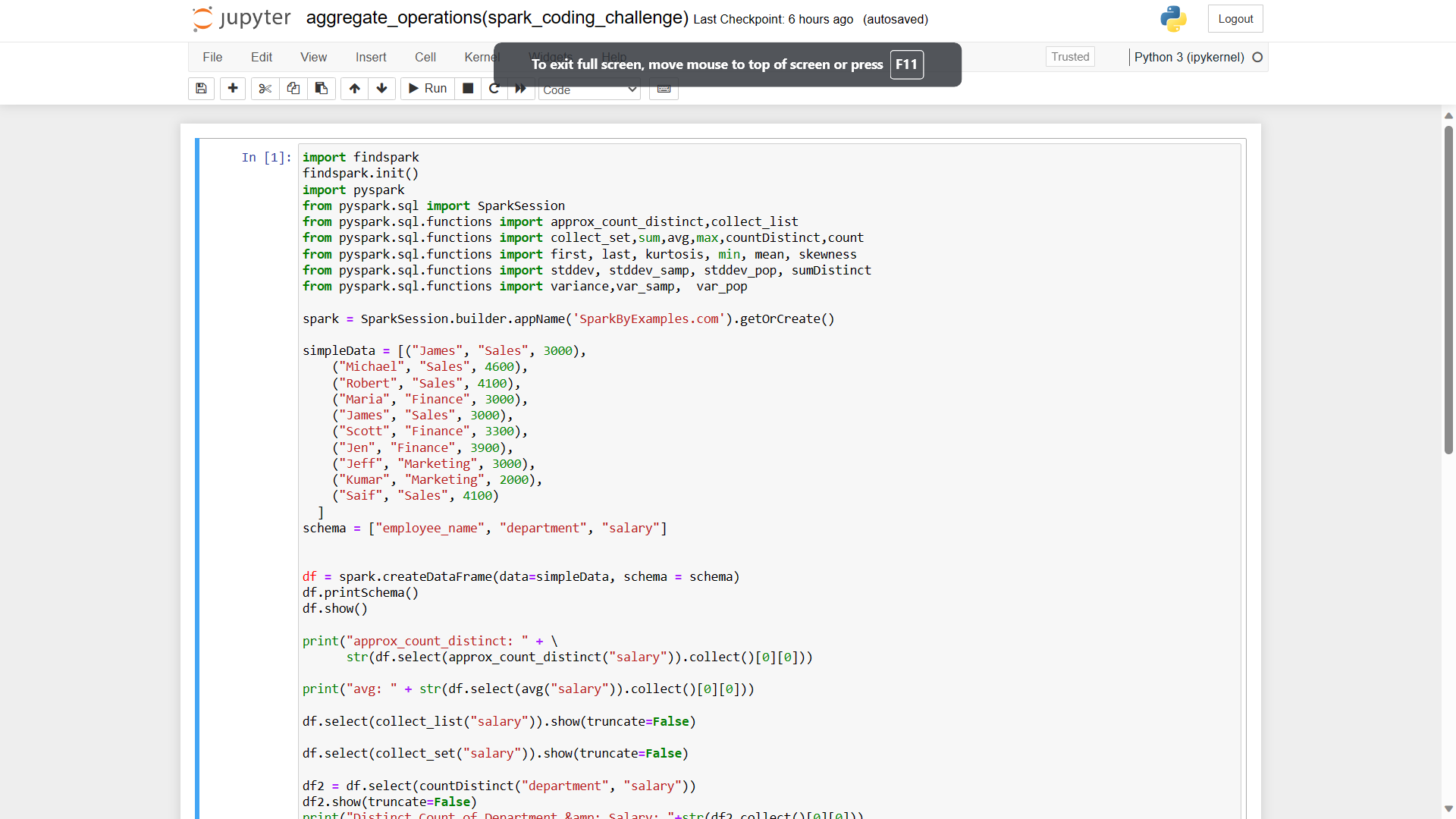


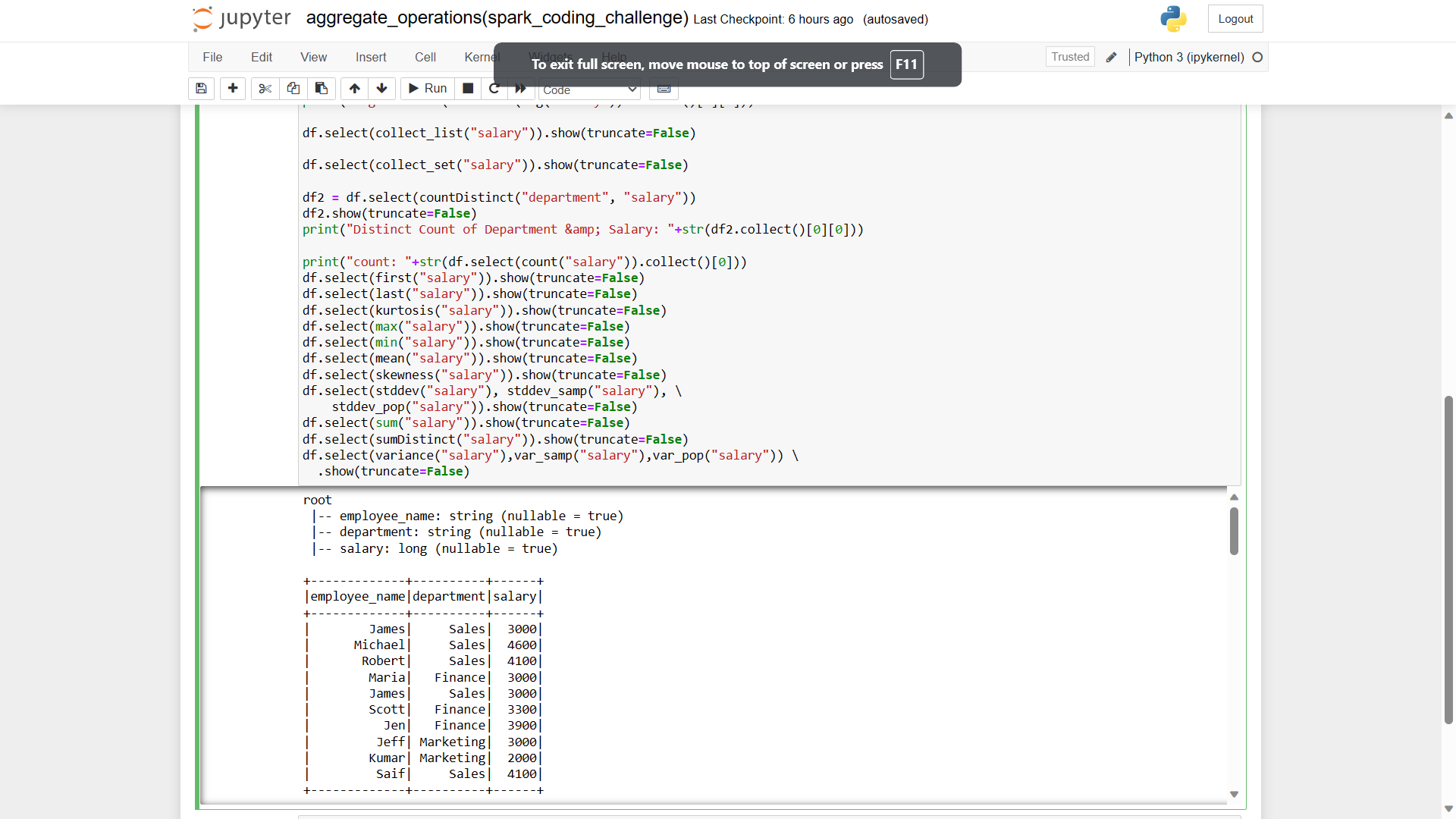
**3) Using Spark SQL - Creating databases, tables**



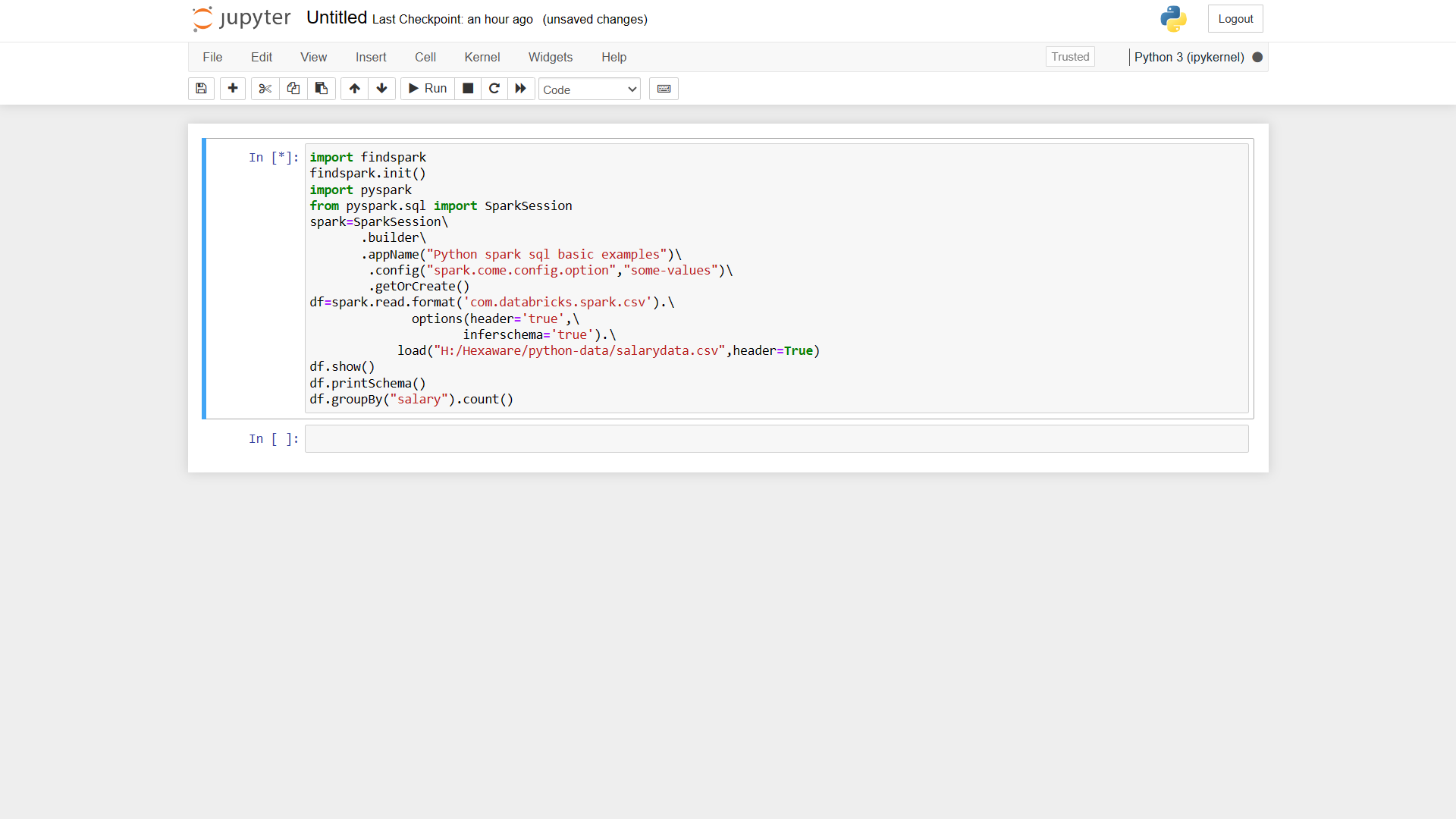
**4) Using Spark SQL - Transformations such as Filter, Join, Simple Aggregations, GroupBy.**

**Aggregate operations:**

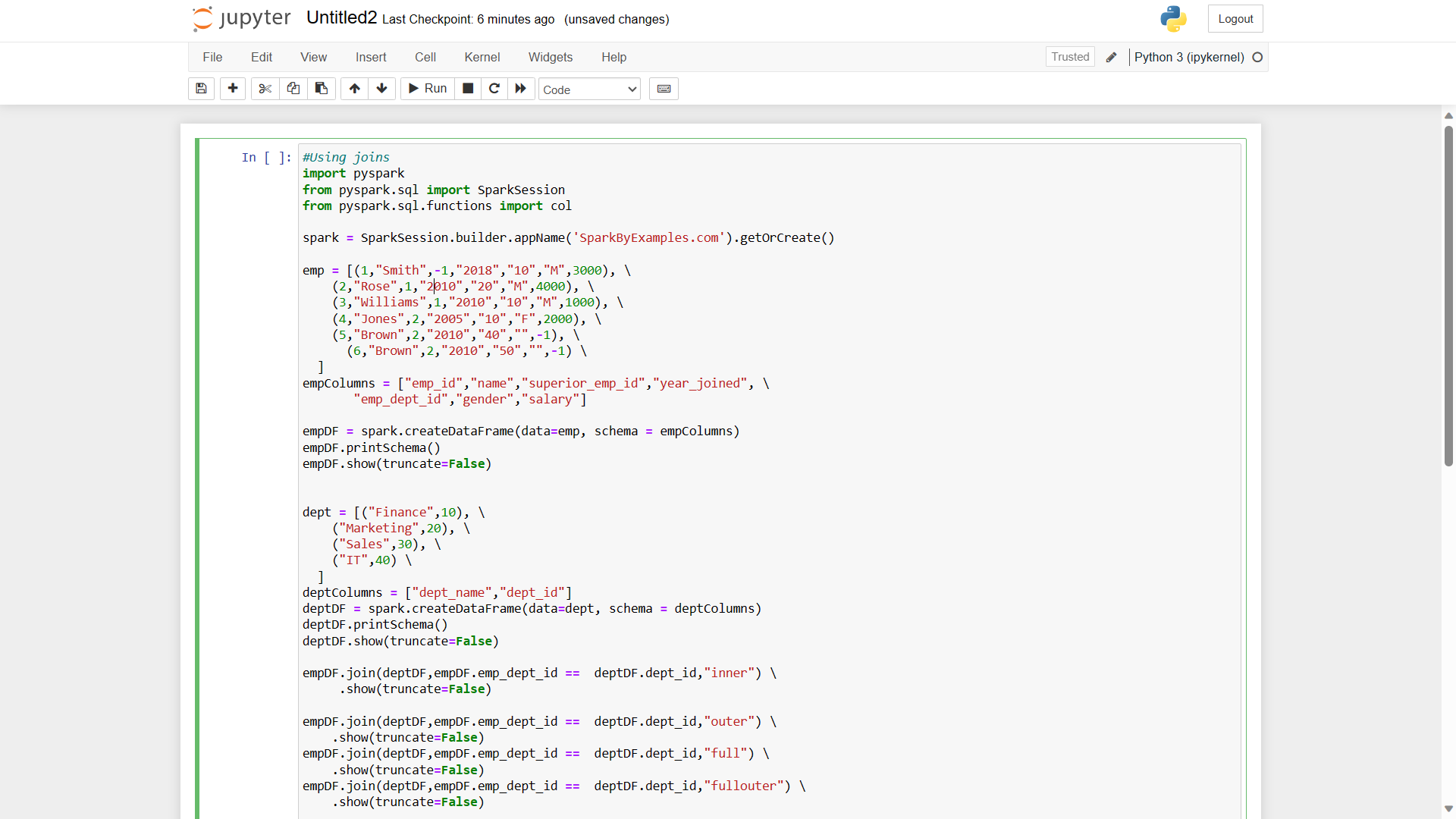


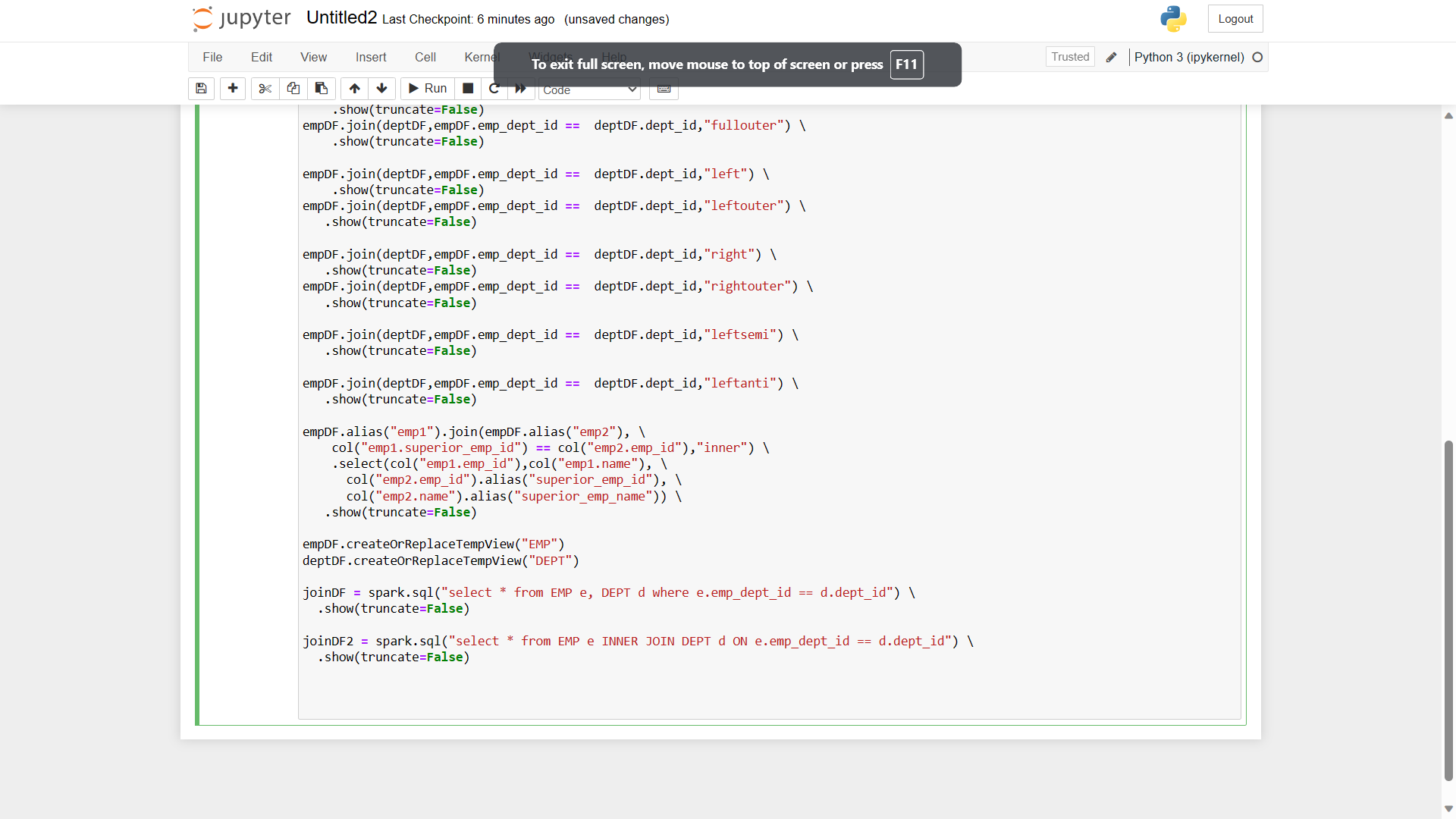


Using group by



Using joins





Using Filters

