Name: Shivprasad CP

Roll No: 9696

Batch: A

Experiment 2

**Implement SQL queries for OLAP operations:**

**Part 1**

**Part 1**

create table *dim\_time*(time\_key *serial* primary key, days *date*, weeks *int*, months *int*, quarter *int*, years *int*);

select \* from dim\_time;

insert into dim\_time(days, weeks, months, quarter, years)

values  ('2016-03-08', 33, 8, 3, 2016),

('2018-06-27', 27, 26, 6, 2018),

('2019-04-30', 30, 18, 4, 2019),

('2015-06-01', 1, 22, 6, 2015),

('2015-12-06', 6, 49, 12, 2015),

('2012-12-21', 21, 51, 12, 2012);

create table *dim\_location*(location\_key *serial* primary key, street *varchar*, city *varchar*, states *varchar*, country *varchar*);

select \* from dim\_location;

insert into dim\_location(street, city, states, country)

values ('gully chowl', 'Navi Mumbai', 'Maharashtra', 'India'),

('Film city road', 'Mumbai', 'Maharashtra', 'India'),

('Sector 24', 'Noida', 'U.P', 'India'),

('Aul market road', 'Patamundi', 'Orrisa', 'India'),

('Shivaji Chowk', 'Pune', 'Maharashtra', 'India');

insert into dim\_location(street, city, states, country)

values ('New market', 'Noida', 'U.P', 'India');

create table *dim\_treatment*(treatment\_key *serial* primary key, treatment\_name *varchar*, costs *int*, medicine\_info *varchar*);

select \* from dim\_treatment;

insert into dim\_treatment(treatment\_name, costs, medicine\_info)

values('Root Canal', 2000, 'Yes'),

('Braces', 2500, 'No'),

('Teeth whitening', 3000, 'Yes'),

('Root Canal', 2500, 'Yes'),

('Wisdom extract', 4000, 'No'),

('Root Canal', 3000, 'Yes');

create table *dim\_diagonse*(diagonse\_key *serial* primary key, diagonse\_info *varchar*, Doctor\_key *int*);

select \* from dim\_diagonse;

insert into dim\_diagonse(diagonse\_info, Doctor\_key)

values('Root Canal', 20),

('Braces', 21),

('Teeth whitening', 40),

('Root Canal', 72),

('Wisdom extract', 34),

('Root Canal', 40);

create table *fact\_revenue*(time\_key *int* references dim\_time(time\_key),

                          location\_key *int* references dim\_location(location\_key),

                          treatment\_key *int* references dim\_treatment(treatment\_key),

                          diagonse\_key *int* references dim\_diagonse(diagonse\_key),

                          net\_revenue *decimal*(19,4), total\_diagonse *int*,

                          primary key(time\_key, location\_key, treatment\_key, diagonse\_key)

                         );

insert into fact\_revenue values(1,1,1,1,20000.50, 10),

(2,2,2,2,10000.50, 5),

(3,3,3,3,15000.50, 8),

(4,4,4,4,22000.50, 12),

(5,5,5,5,24000.50, 15),

(6,6,6,6,18000.50, 9);

select \* from fact\_revenue;

drop table fact\_revenue;

*-- Roll Up Operations*

select city, states, sum(net\_revenue) from dim\_location inner join

fact\_revenue on dim\_location.location\_key *=* fact\_revenue.location\_key

group by rollup(states, city) order by states,city;

*--Cube operations*

select years, quarter, total\_diagonse, sum(net\_revenue) from fact\_revenue natural inner join dim\_time

group by cube(years, quarter, total\_diagonse);

*--Slice Operations*

select states, city, sum(net\_revenue) from dim\_location inner join fact\_revenue on

fact\_revenue.location\_key *=* dim\_location.location\_key where states *=* 'Maharashtra' group by states, city;

*--Dice Operations*

select states, city, sum(net\_revenue) from dim\_location inner join fact\_revenue on

fact\_revenue.location\_key *=* dim\_location.location\_key where states *=* 'Maharashtra' and city *=* 'Mumbai'

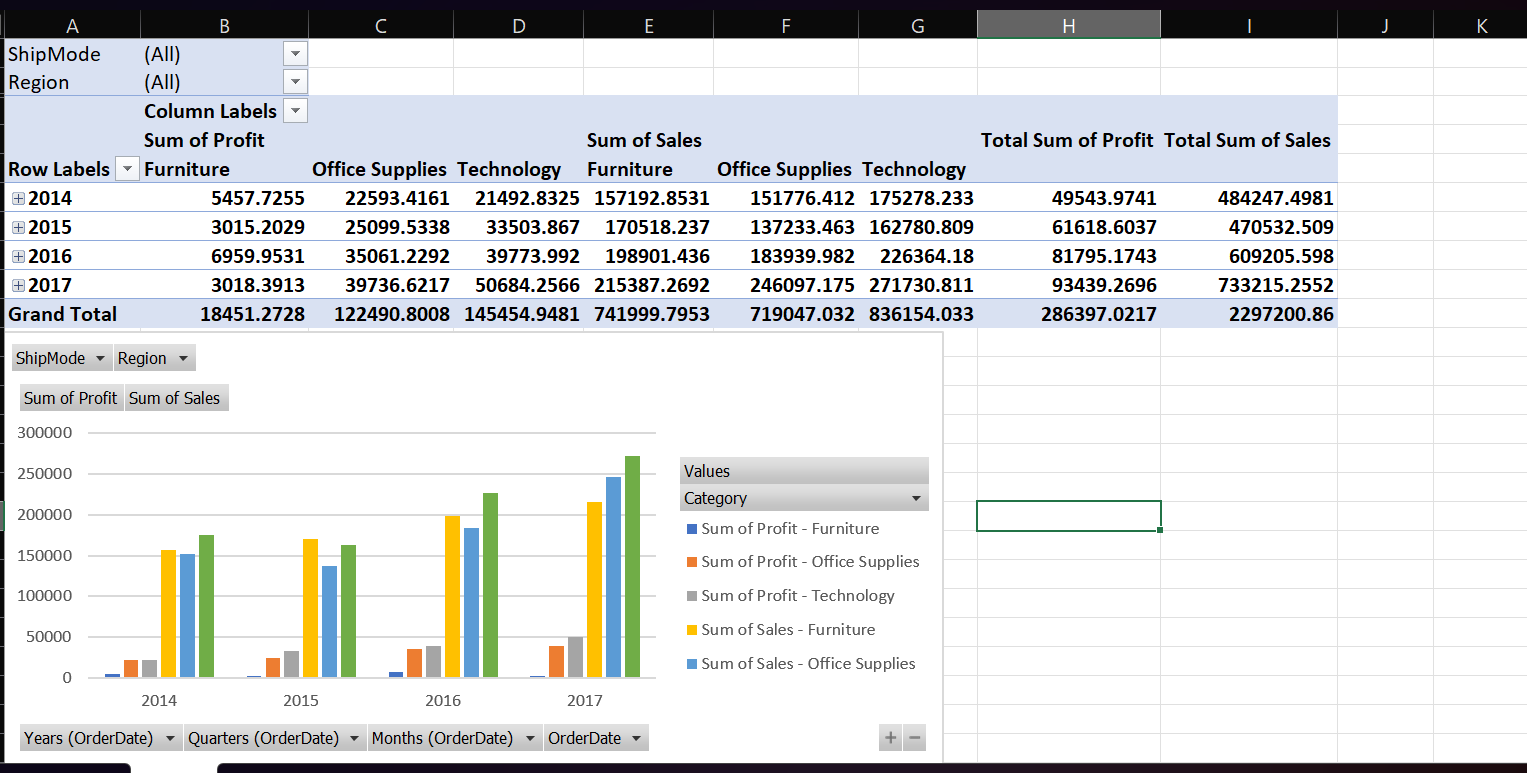
group by states, city;

Part 2:

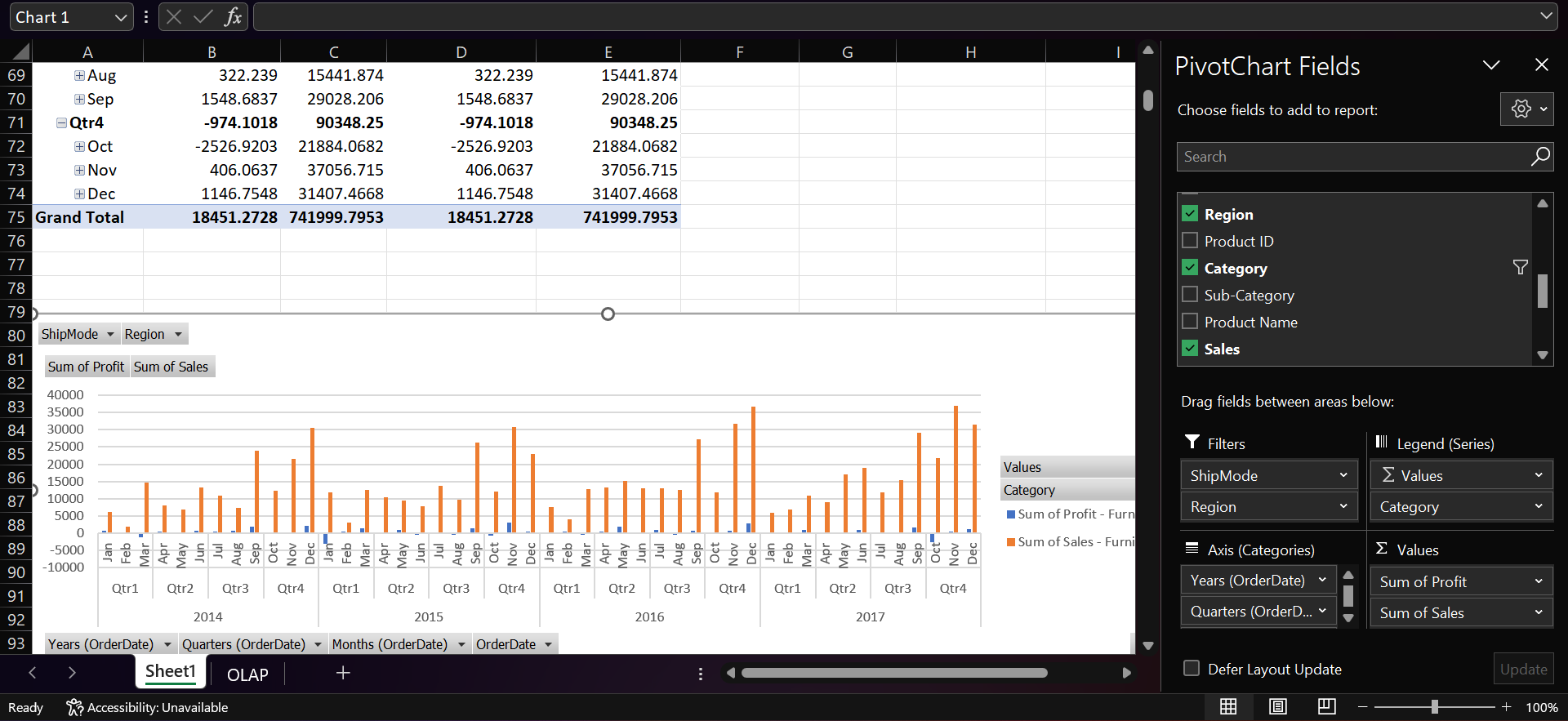
Perform data analysis with visualization for the following:

a)To view monthly, quarterly, yearly profit, sales of each category, region wise

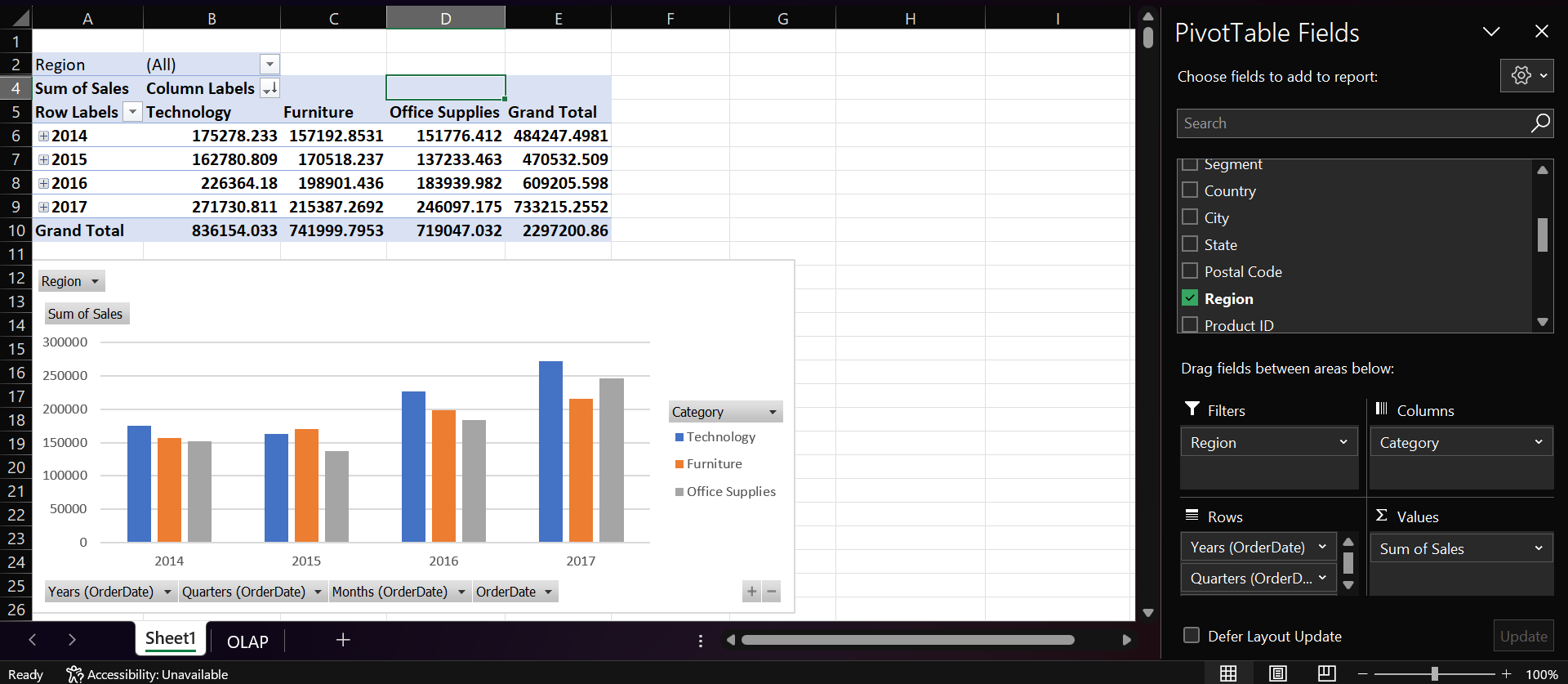
b) Comparison of sales and profit on various years.



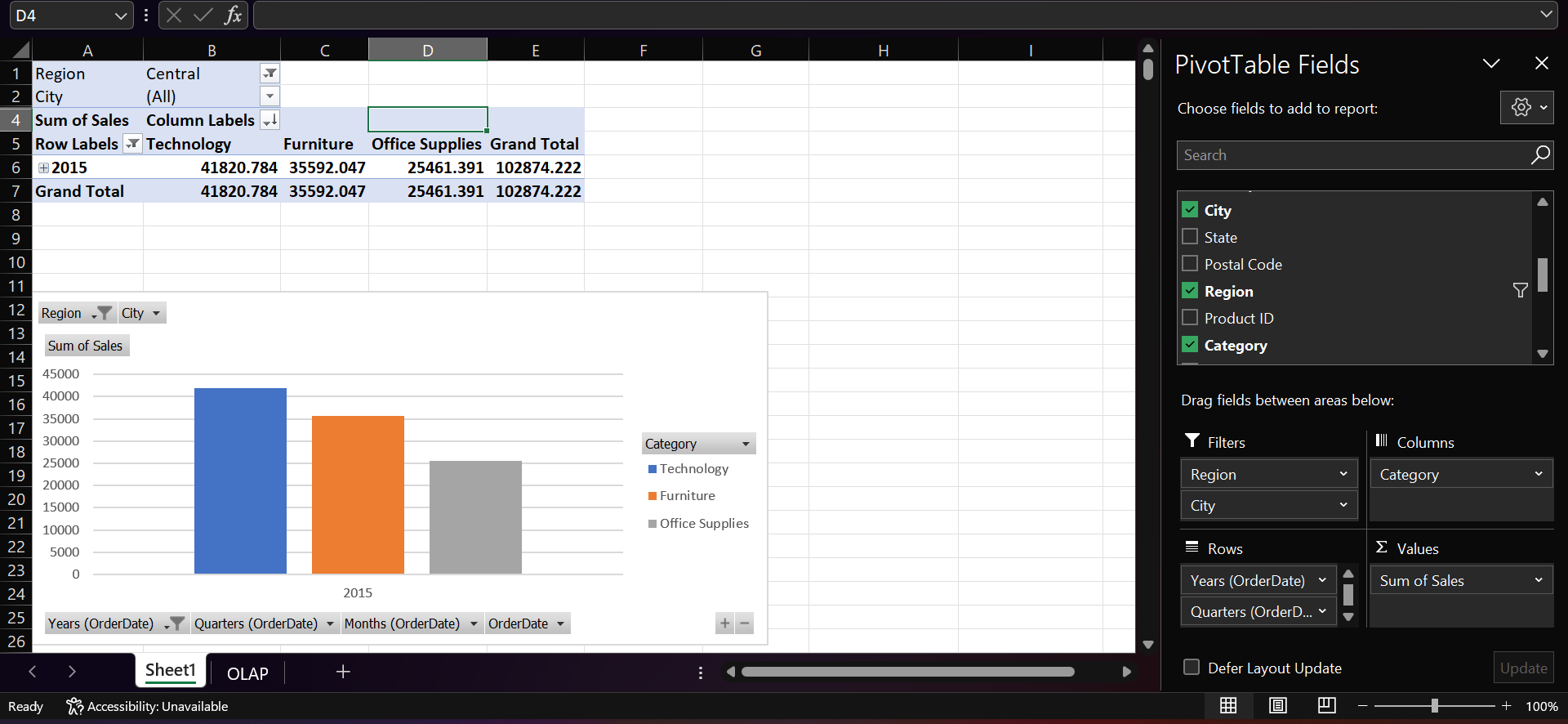
c) Comparison of sales in various months for product category =furniture.



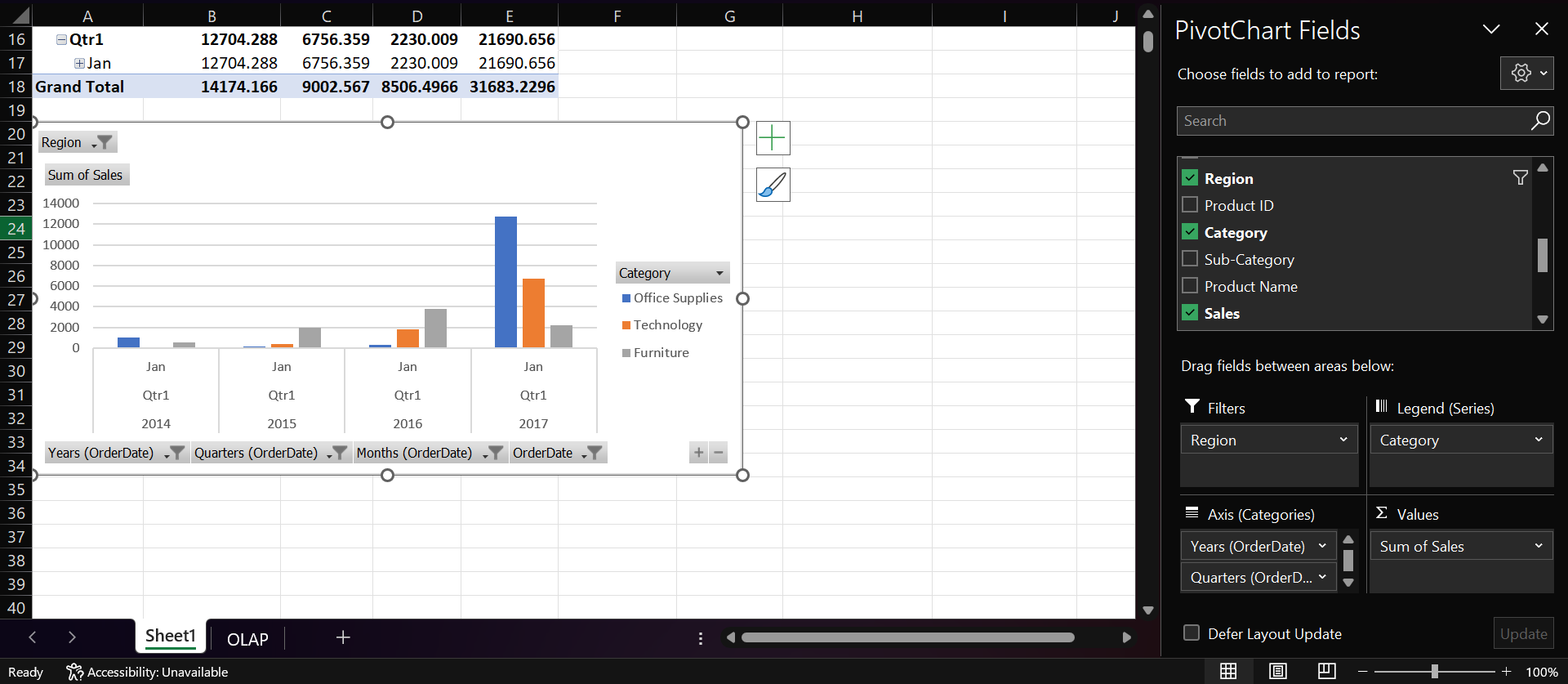
d) Need to know which product has more demand on which location?



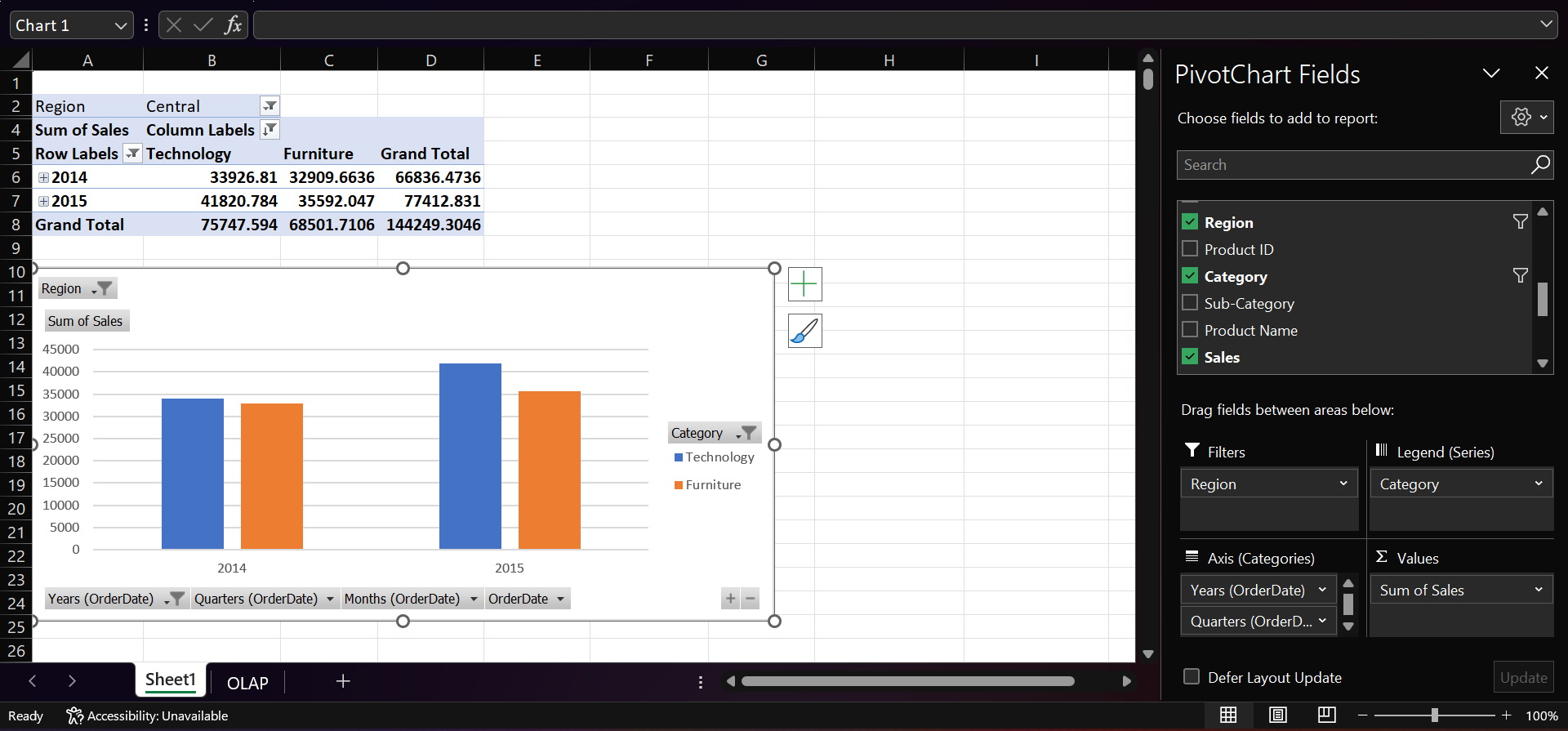
e) What is sale of product category wise, city wise for year=2015 ?



f) What is sales of product categories of Jan month of each year?



g) What is trend of sales on year 2014 and 2015 for product category furniture and technology?



Part 3:

List down the data Extraction Transformation and Loading processes applicable for your DW system.

Extract:

* Data Source Identification: Different sources of dental clinical data, which could include electronic health records, appointment systems, billing systems, patient information, etc.
* Data Extraction: Extract relevant data from the identified sources using appropriate methods, such as database queries, flat file exports, etc.
* Data Profiling: Analyze and profile the extracted data to understand its quality, structure, and potential issues.

Transform:

* Data Cleansing: Cleaning the extracted data by identifying missing values, inconsistencies, errors, and duplicates.
* Data Transformation: Convert data into a common format and standardize units, terminologies, and codes to ensure consistency.
* Data Integration: Integrate data from various sources into a unified format, considering data types, relationships, and hierarchies.
* Data Validation: Validate the transformed data to ensure that it meets the defined quality standards and business rules.

Loading:

* Staging: Store the cleaned and transformed data in a staging area, separate from the data warehouse, to facilitate further validation.
* Data Warehouse Loading: Load the validated and transformed data into the data warehouse.