**What is Data ?**

Data is nothing but systematically recorded values and facts about a quantity. When the data available to us is not systematic or Organized, they are known as Raw Data. Mostly, the data given to us is in form of Raw data, and systematically Organizing them may be in form of either Bar Graph, Pictograph, Double Bar graph, or any other form of visual representation is called as Organization of Raw Data.

| **Data** | **Information** |
| --- | --- |
| Data is the raw fact. | It is a processed form of data. |
| It is not significant to a business. | It is significant to a business. |
| Data is an atomic level piece of information. | It is a collection of data. |
| Example: Product name, Name of student. | Example: Report card of student. |
| It is a phenomenal fact. | It is organized data. |
| This is the primary level of intelligence. | It is a secondary level of intelligence. |
| May or may not be meaningful. | Always meaningful. |
| Understanding is difficult. | Understanding is easy. |

**Why it is important ?**

The role of data is changing rapidly, helping organizations to actively transform, providing the following:

1. **Operational efficiency**: reducing costs, reducing heavy manpower wherever applicable, concentrating and directing resources where required, unforeseen expenditure, resulting in business model being more efficient and competitive.

2. **Understanding customer**: by providing crucial information, the marketing and sales teams can plan their strategies according to the customer requirements and buying patterns to provide a personalized customer experience and thus, increase sales and customer retention.

3. **Risk management**: the pandemic has taught us that the future is uncertain. Change can happen very rapidly. The information from data helps businesses to plan and be prepared with better-informed decisions and strategies.

4. **Business expansion**: data helps understand consumer behaviour and helps companies to learn how to harvest, use and monetize their data. Add to the product range and bring positive changes to the existing ones.

5. **Personalized experience**: The mobile applications let companies gather multiple data points on consumer, such as clients’ geolocation, the most common user interactions, user behaviour and browsing history. This information then be utilized to make up for a lack of face-to-face connection with clients.

**What is Data organisation and management?**

Data organization is the way to arrange the raw data in an understandable order. Organizing data include classification, frequency distribution table, picture representation, graphical representation, etc.

Data organization helps us to arrange the data in order that we can easily read and work. It is difficult to work or do any analyses on raw data. Hence, we need to organize the data to represent them in a proper way. Let us understand with the help of an example.

**Data management** is the practice of collecting, keeping, and using data securely, efficiently, and cost-effectively. The goal of data management is to help people, organizations, and connected things optimize the use of data within the bounds of policy and regulation so that they can make decisions and take actions that maximize the benefit to the organization. A robust [data management strategy](https://www.oracle.com/in/database/) is becoming more important than ever as organizations increasingly rely on intangible assets to create value.

**What is Data Scalability ?**

Database scalability refers to the ability of a database to handle increasing amounts of data, numbers of users, and types of requests without sacrificing performance or availability. A scalable database tackles these database server challenges and adapts to growing demands by either adding resources such as hardware or software, by optimizing its design and configuration, or by undertaking some combined strategy.

Vertical database scalability means adding more processing power and memory to a single server. Horizontal database scalability means adding more servers (i.e., database nodes) to distribute the workload.

Techniques such as sharding or replication are integral to horizontal scalability. These methods partition and distribute data across multiple servers to improve performance and reliability.

Database scalability is important in applications that need (or anticipate the need) to handle large volumes of data, such as e-commerce, social media, and finance. A scalable database design can help businesses avoid bottlenecks, reduce downtime, and ensure a positive user experience even as demand increases.

**What is RDBMS ?**

A relational database is a type of database that stores and provides access to data points that are related to one another. Relational [databases](https://www.oracle.com/in/database/what-is-database/) are based on the relational model, an intuitive, straightforward way of representing data in tables. In a relational database, each row in the table is a record with a unique ID called the key. The columns of the table hold attributes of the data, and each record usually has a value for each attribute, making it easy to establish the relationships among data points.

**What are the important features of RDBMS ?**

Here are some of the basic features of a relational database management system.

* It offers functions related to creating, reading, updating, and deleting operations to be performed on data.
* It provides data dictionaries and metadata collections which help in data handling.
* Another common feature is data storage management. The data objects range from binary large objects to strings to stored procedures.
* SQL is the most commonly used language to access RDBMS. With certain extensions, it can be used with common programming languages such as Common Business-oriented Language, Java, etc.
* Multiple users can concurrently access a database without the data integrity being compromised.
* It provides security management since it is used mostly in enterprises.
* It supports the work of database administrators who must manage and monitor database activity. With the help of utilities, data loading, and database backup are automated.
* It manages log files that track system performance based on selected operational parameters.
* It provides a graphical interface that helps database administrators visualize database activity.

**What is Data Consistency ?**

* Consistency means having the property of always maintaining the same standards, opinions, and actions. In order to keep the database consistent, data written to the database must be valid according to all defined rules.
* Data Consistency in DBMS is defined by a set of rules that ensure that all data points in the database system are correctly read and accepted. This is achieved by making rules. Transactions of data written to the database should only modify the affected data established by rules set by the database developer. It doesn't guarantee that the transaction is correct, it just guarantees that it doesn't violate the rules defined by the program.

**Constraints**

* [NOT NULL](https://www.w3schools.com/sql/sql_notnull.asp) - Ensures that a column cannot have a NULL value
* [UNIQUE](https://www.w3schools.com/sql/sql_unique.asp) - Ensures that all values in a column are different
* [PRIMARY KEY](https://www.w3schools.com/sql/sql_primarykey.asp) - A combination of a NOT NULL and UNIQUE. Uniquely identifies each row in a table
* [FOREIGN KEY](https://www.w3schools.com/sql/sql_foreignkey.asp) - Prevents actions that would destroy links between tables
* [CHECK](https://www.w3schools.com/sql/sql_check.asp) - Ensures that the values in a column satisfies a specific condition
* [DEFAULT](https://www.w3schools.com/sql/sql_default.asp) - Sets a default value for a column if no value is specified
* [CREATE INDEX](https://www.w3schools.com/sql/sql_create_index.asp) - Used to create and retrieve data from the database very quickly

**Advantages of RDBMS**

➨It is easy to use.  
➨It is secured in nature.  
➨The data manipulation can be done.  
➨It limits redundancy and replication of the data.  
➨It offers better data integrity.  
➨It provides better physical data independence.  
➨It offers logical database independence i.e. data can be viewed in different ways by the different users.  
➨It provides better backup and recovery procedures.  
➨It provides multiple interfaces.  
➨Multiple users can access the database which is not possible in DBMS.

**Disadvantages of RDBMS**

➨Software is expensive.  
➨Complex software refers to expensive hardware and hence increases overall cost to avail the RDBMS service.  
➨It requires skilled human resources to implement.  
➨Certain applications are slow in processing.  
➨It is difficult to recover the lost data.

**Acid Properties**

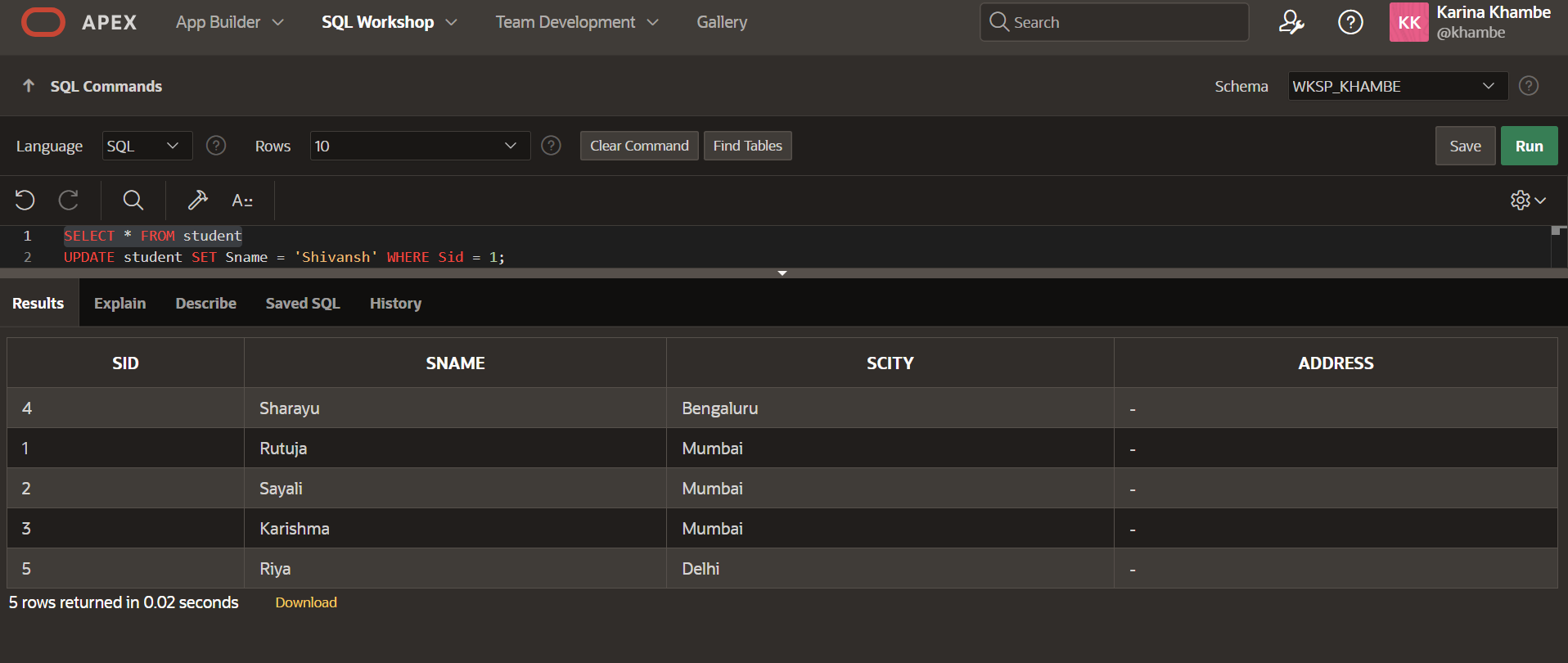
ACID is an acronym that refers to the set of 4 key properties that define a transaction: Atomicity, Consistency, Isolation, and Durability. If a database operation has these ACID properties, it can be called an ACID transaction, and data storage systems that apply these operations are called transactional systems. ACID transactions guarantee that each read, write, or modification of a table has the following properties:

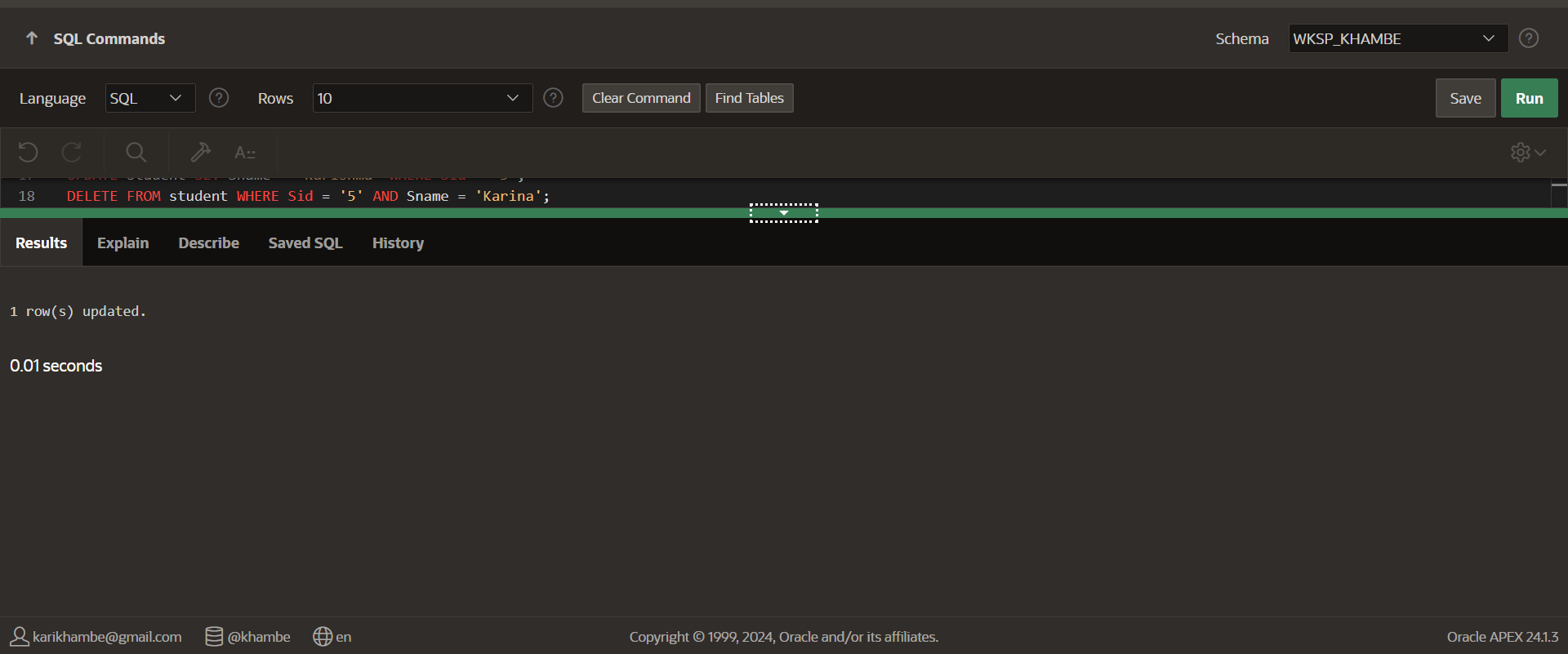
• Atomicity - each statement in a transaction (to read, write, update or delete data) is treated as a single unit. Either the entire statement is executed, or none of it is executed. This property prevents data loss and corruption from occurring if, for example, if your streaming data source fails mid-stream.

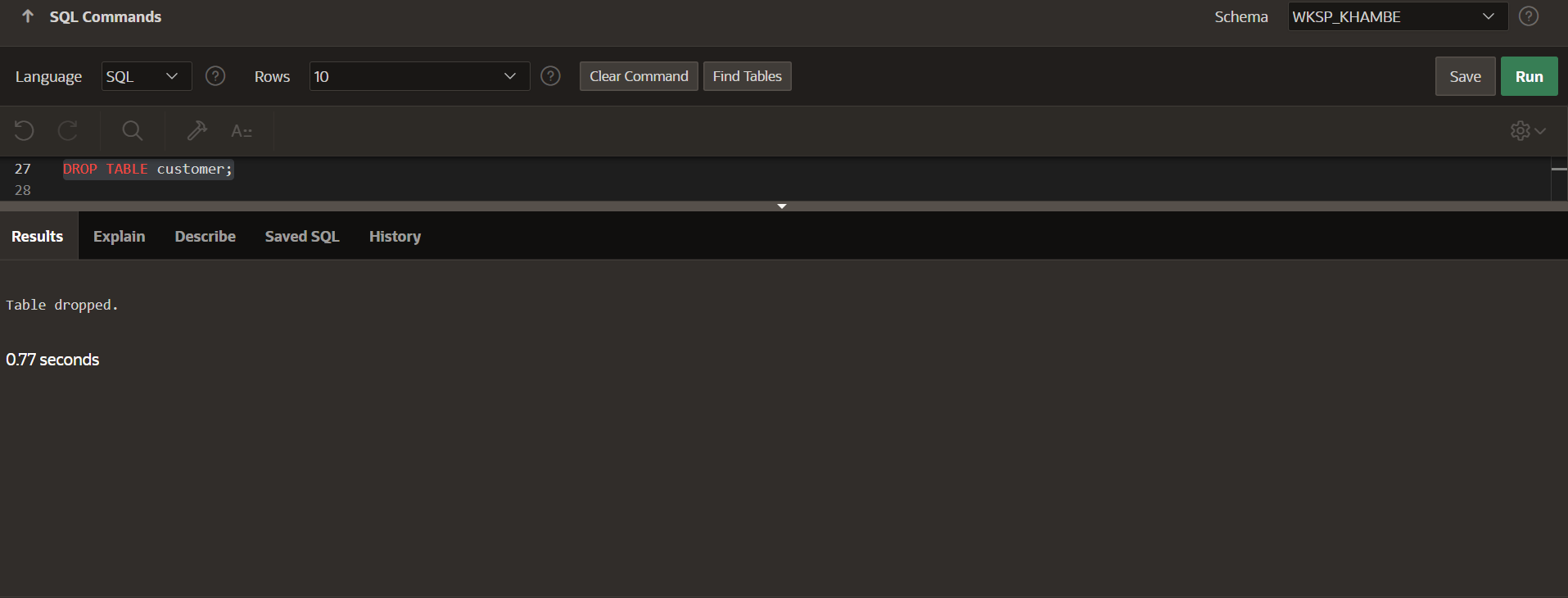
• Consistency - ensures that transactions only make changes to tables in predefined, predictable ways. Transactional consistency ensures that corruption or errors in your data do not create unintended consequences for the integrity of your table.

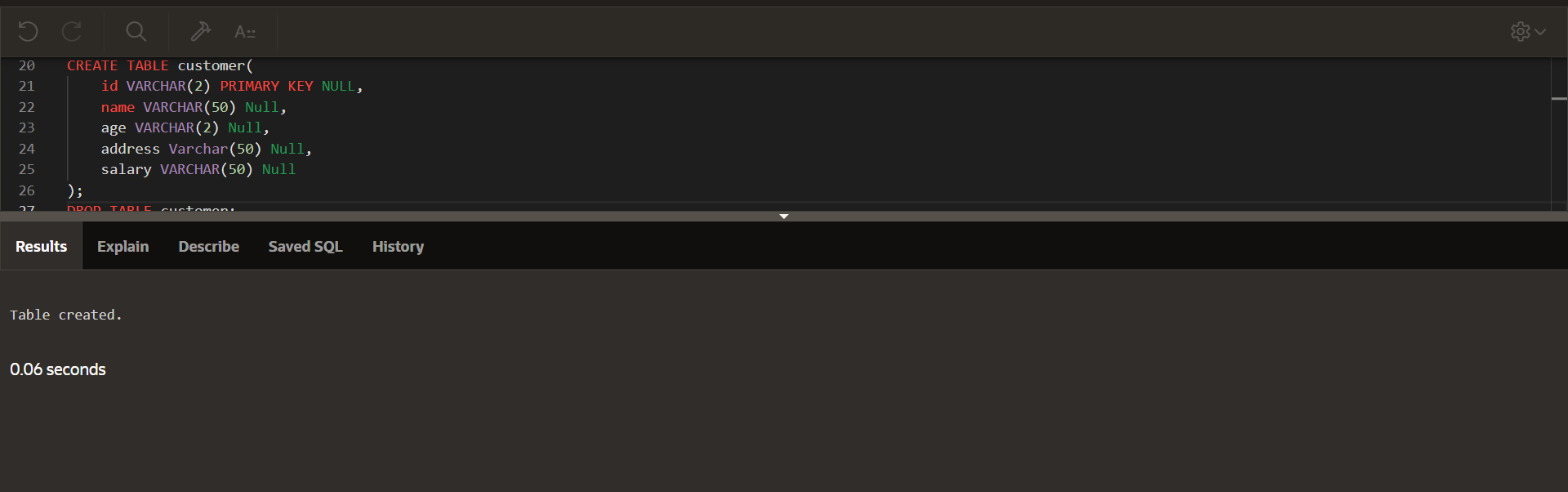
• Isolation - when multiple users are reading and writing from the same table all at once, isolation of their transactions ensures that the concurrent transactions don't interfere with or affect one another. Each request can occur as though they were occurring one by one, even though they're actually occurring simultaneously.

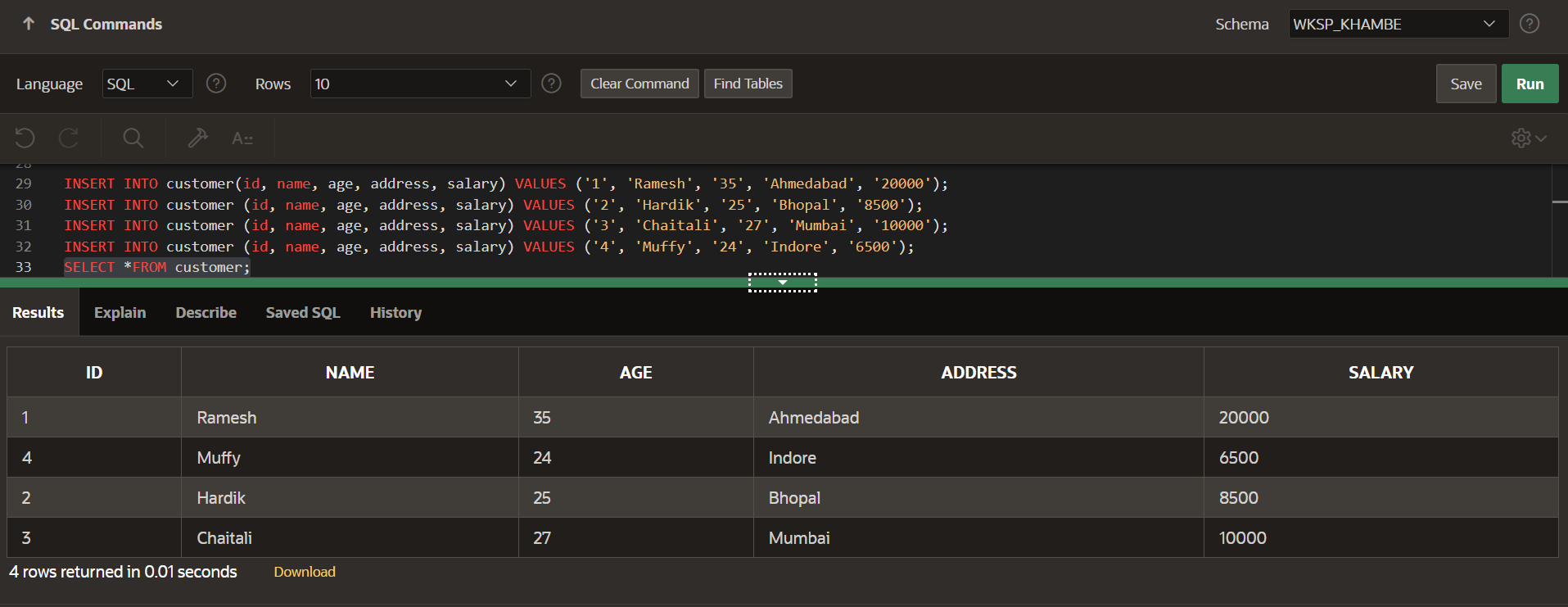
• Durability - ensures that changes to your data made by successfully executed transactions will be saved, even in the event of system failure.

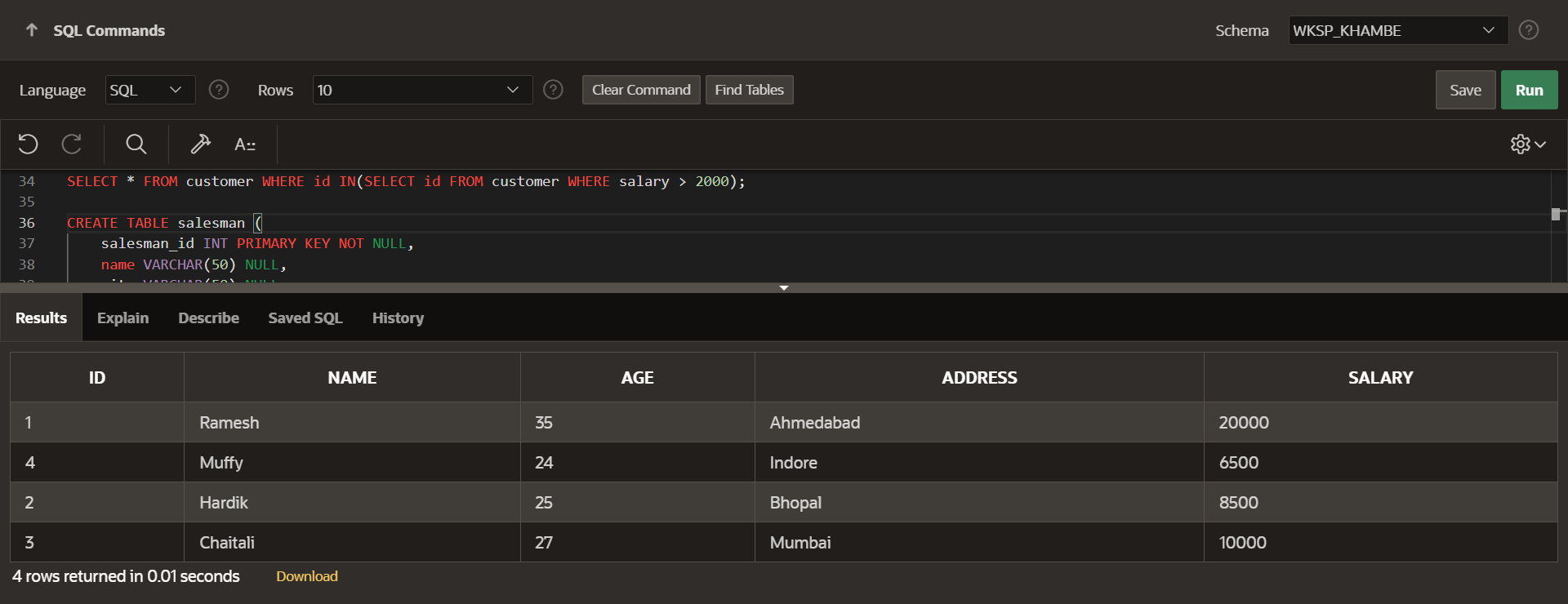
****

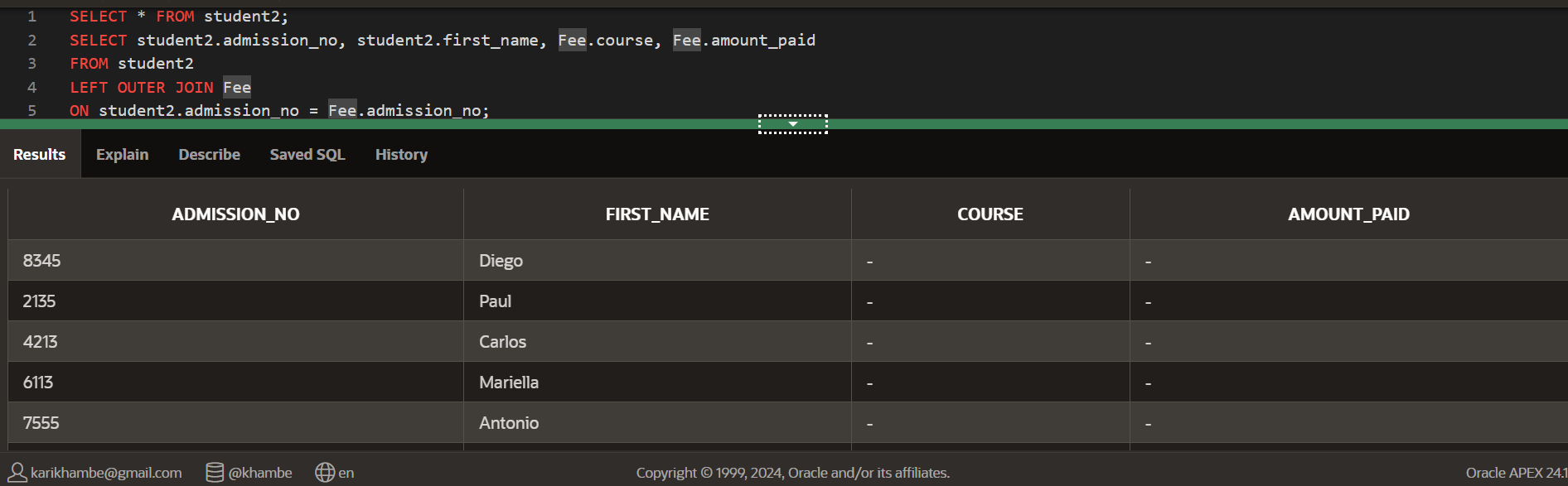
****

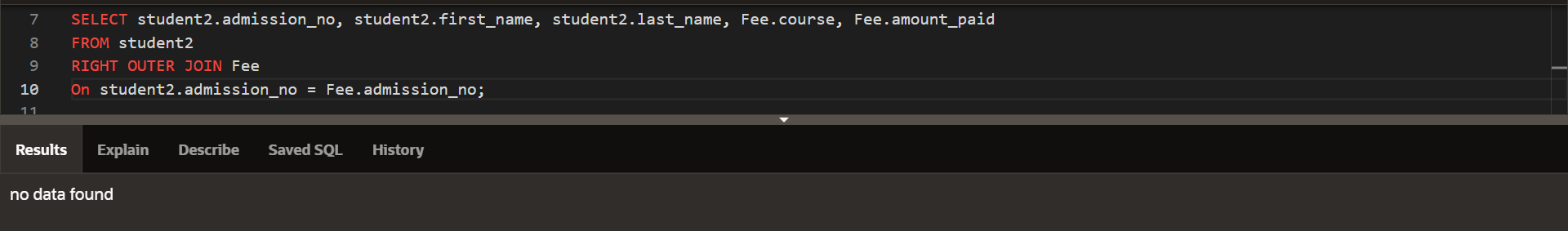
****

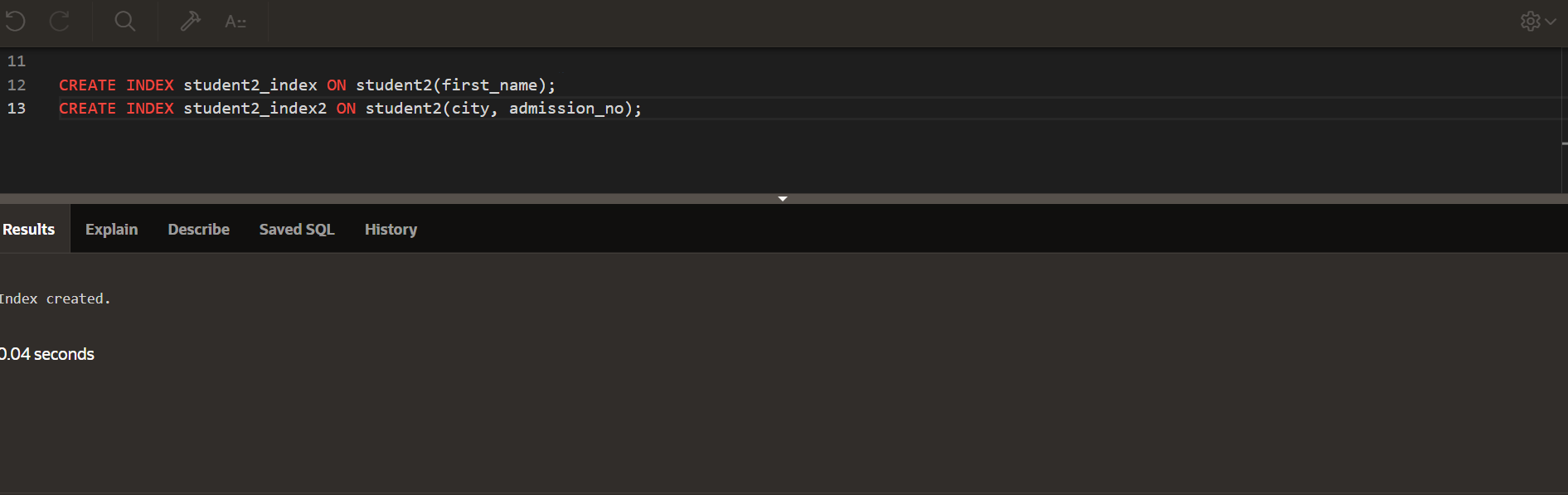
****

****

****

****

****

****

**ACID is an acronym that refers to the set of 4 key properties that define a transaction: Atomicity, Consistency, Isolation, and Durability. If a database operation has these ACID properties, it can be called an ACID transaction, and data storage systems that apply these operations are called transactional systems. ACID transactions guarantee that each read, write, or modification of a table has the following properties:**

**• Atomicity - each statement in a transaction (to read, write, update or delete data) is treated as a single unit. Either the entire statement is executed, or none of it is executed. This property prevents data loss and corruption from occurring if, for example, if your streaming data source fails mid-stream.**

**• Consistency - ensures that transactions only make changes to tables in predefined, predictable ways. Transactional consistency ensures that corruption or errors in your data do not create unintended consequences for the integrity of your table.**

**• Isolation - when multiple users are reading and writing from the same table all at once, isolation of their transactions ensures that the concurrent transactions don't interfere with or affect one another. Each request can occur as though they were occurring one by one, even though they're actually occurring simultaneously.**

**• Durability - ensures that changes to your data made by successfully executed transactions will be saved, even in the event of system failure.**