Unit 3.
Reporting Authoring.

* Relational Data Model.

- type of data organization used in relational database
- represents data in the form of tables.
- tables contain rows & columns.
- tables are structured in such a way that relationships can be established betⁿ them using keys (shared)
- each table → entity. (rep) → customers / products / orders.
- row → instance of entity.
- column → attribute / characteristic of entity.
- Eg → A customer table may have columns for Customer_ID, name, address etc
- relationships betⁿ tables → linking data from one table to another table via shared keys.
- keys are unique identifiers → primary keys & foreign keys.
- Relational Data allows efficient storage, retrieval and manipulation of data through SQL!
- provide flexible & organized way to store & manage large amount of data.

* How to build Reports with Relational Data Model.

Steps:

1) Identify the report requirements
- Determine specific information / insights you want to present in the report.
- Understand data elements, metrics & relationships between the tables that are relevant.

2) Design the query.
compose an SQL query that retrieves the necessary data from the relational database

3) Execute the query.
Run the SQL query against the database to fetch the required data.

4) Process & Format the Data.
Process the retrieved data & transform it to a format suitable for report. (data formatting / manipulation → presentation)

5) Choose reporting tool / platform
- Choose reporting tool that suits your requirements and preferences
- Available options → Tableau, Power BI etc.

6) Select rep

6) Design report layout
- design visual layout of your report (reporting tool)
- arrange data in tables, graphs, charts based on sq.

7) Bind data to the report
- connect the processed data from query to report.
- Associate data fields with corresponding report elements such as tables, charts or labels.

8) Customize the report.
- Enhance report by formatting, styling, branding elements etc
- consider adding filters, parameters, etc.

9) Generate & distribute report
- use the reporting tool to generate report in desired format such as Pdf, Word etc

10) Maintain the / update the report
- Regularly review & update the report as needed.
- Monitor the data quality, refresh the report data to reflect the latest information.

Advantages

1) It is easy to understand & navigate, making it accessible to users with various levels of technical expertise.
2) It enforces data integrity through primary key & foreign key relationships.
help to maintain the consistency & accuracy.
3) It is highly flexible & scalable.
4) Provide powerful query languages such as SQL for retrieval manipulation & analysis of data.

Disadvantages.

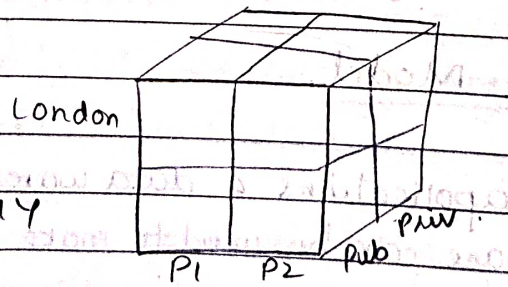
1) Relational database face performance issues when dealing with complex queries.
2) Scalability issues when scaling horizontally across multiple servers.
3) As it is defined by a schema there is limited flexibility.

Multidimensional Model

- best for OLAP applications & data warehouses
- relational database can be used to make multidimensional model.
- contains arrays of 3 or more dimensions.
- In 2-D, we have X & Y, in multi, we have X, Y, Z etc depending on the no. of dimensions.
- used for OLAP applications & data warehousing.
- represents data in a cube form
- called data cube.
- Data cube enables data to be modeled & viewed in various dimensions.

How to build a Dimensional Data Model?

- Dimensional data model consists of dimension tables & fact tables
- fact table: Stores numeric information about different business measures.
- Dimension tables: store attributes used to describe objects in the fact table
- Primary & foreign keys connect fact tables & dimensions
- ⁂ can be built based on various schemas such as star, snowflake or galaxy.
- The centre of every star schema has a fact table.
- This fact table contains measures & foreign keys of connected dimensions.
- ordering data in database with good arrangement & assembling contents in database.
- allows user to ask analytical questions related to market & business trends unlike Relational model (queries)
- ,



- 3rd dimension: Customer Type < pub. private.
- 4th dimension can be time.

Advantages.

1) Grouping of similar information effectively.
2) Easy visualization of complex data.

3) Easily identify patterns, threads-trends, outliers and anomalies from data.
4) Extract valuable information from unstructured data.
5) Suitable for both, structured & unstructured data.
6) Processes data quickly.
7) Easy maintainance of data.

Disadvantages.

1) It has limited flexibility.
2) It has difficulty in handling sparse data.
3) It involves data redundancy.
4) It focuses on grouping data along dimensions therefore cannot represent complex representations & hierarchies.

Types of Reports.

1) List Report

- basic form of report
- base of reporting about data & its insights.
- presents data in a tabular format
- displays a list of records/items.
- straigtforward representation of info. without complex analysis or visualisation.
- presents data in a table or grid format with rows & columns.
- Each row → record/item Column → attribute/field of data
- often provide options for sorting & grouping the data based on specific attributes.
- Sorting arranges data rows in ascending order/des. order

- Grouping organizes data rows into logical groups based on shared attributes.
- In cases where data exceeds available space on a single page, list report includes pagination
- Pagination breaks the report into multiple pages, allowing users to navigate through data.
- often include headers & footers that provide additional info. about the report, such as report title, date, page nos etc
- offer options to export data in different formats such as Pdf or excel or print it directly.

2) Crosstab Report.

- also known as pivot table report.
- represents data in a matrix-like format.
- provides summarized view of relationships between two or more variables.
- allows users to analyze data across different dimensions
- it provides insights into intersection of data categories.
- data is organized in matrix format where row & columns represent different categories of data.
- ~~Each cell~~ Row categories represent one dimension of data & typically displayed along vertical axis
- Row categories are often hierarchial & can represent factors such as time periods, product categories etc.
- Column categories represent another dimension of data & displayed along horizontal axis.
- Cells with matrix display aggregated values based on combination of row & column categories.
- These values can include sum, avg, % etc. depending on data & its analysis.

- Provide drill-up & drill down capabilities.
- users can drill up to view higher level summaries & drill down to explore more granular data.
- may also include data visualisation such as color coding, heatmaps, bar charts etc to highlight patterns & trends.

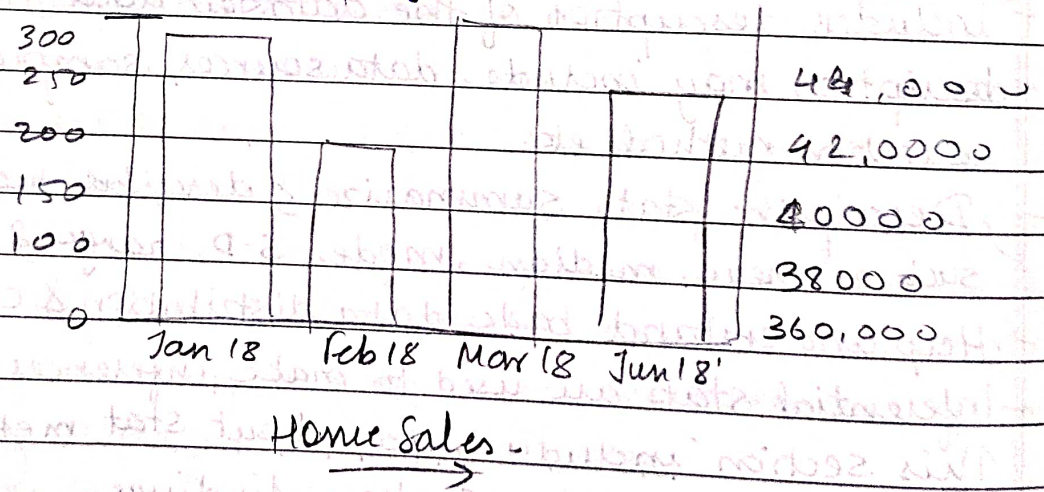
3) Statistical Report.

- report that presents findings based on statistical operations applied to the dataset
- provides comprehensive ~~view~~ overview of data including descriptive statistics, inferential stats & interpretation of results.
- starts with an introduction that provides background info about study / research.
- data description
- includes description of the dataset used for analysis.
- description may include data sources, sample size, data collection methods etc
- Descriptive stats summarize & describe main characteristics such as mean, median, mode, S.D, range & percentiles.
- Help understand basic data distribution & central tendencies
- Inferential stats are used to make inferences & draw conclusion
- This section includes info about stat methods used.
- Result section presents key findings.

4) Chart Report

- type of report that presents data using visual charts or graphs.
- chart - graphic that displays numeric data in a compact visual layout & reveals essential data relationships.

- common types of charts include line charts, bar charts, pie charts, scatter plots etc.
- Choice of chart depends on nature of data & specific message or insight you want to communicate.
- The main focus of chart report is data visualisation.
- Visual elements such as colors, labels, legends & scales should be carefully chosen to convey intended message.
- The scales, axes & data labels should be appropriately labeled & annotated to provide context
- involve analyzing trends over time or comparing variables.
- Include clear & informative data labels.
- typically includes introduction describing purpose & scope of analysis, charts/graphs representing data, captions/explanation for each chart & conclusion that summarizes the key insights or findings.



5) Map Report

- visually uses maps to represent geographic data & information.
- combines spatial data with textual/numeric data to provide insights into geographical patterns, distribution & relationships
- commonly used in fields such as geography, urban planning etc

- often incorporates multiple layers of data such as boundaries, points of interest, etc.
- may be accompanied by tables, charts or graphs that provide add. info.

6) Financial report

- comprehensive document that provides overview of organizat ion's financial performance & position.
- summarizes financial data including income, expenses, assets, liabilities, equitys presents it in a structured format.
- essential for stakeholders to assess financial health & stability of an organization.
- key components : balance sheet, income statement, cashflow statement, statement of shareholder's equity etc.
- balance sheet - organization's assets, liabilities & shareholders equity at a specific point.
- Income statement - reports org. revenue, profit, loss over a specific period.
- Cashflow statement - reports et cash inflow & outflows during specific period.

Operations on Data Reports.

Data Grouping.

- categorizing / organizing data based on specific crite or attributes.
- involves grouping related data together to facilitate analysi presentation & intepretation of information in a structured

manner.
- By grouping data you can identify patterns, trends & relationships in dataset easily.
- involves selecting a field or group attribute that you want to group the data by.
- Eg. if we have a sales report with info about products, customers & sales figures, you may choose to group data by product category or customer location.
- So grouped data can then be aggregated, summarized & presented in various ways through charts, tables or graphs to provide clearer understanding of the information.
- allows you to perform calculations, comparisons & analysis within each group, enabling you to make data-driven decisions
- useful in reports dealing with huge datasets or complex information.
- helps in organizing data to enhance readability.
- enables efficient analysis of data.

Filtering Reports.

- process of analyzing data & extracting relevant information based on specific criteria or parameters
- ~~It~~ You involves setting up a certain conditions or rules to narrow down data & focus on specific aspects of it interest.
- conditions can be based on various factors such as time-frames, geographical locations, specific categories etc.
- Filters help in retrieving specific data that is relevant/essential for making decisions.

Eg. sales report can be filtered to show data for a particular region / product category.

- Filtering reports allows analysts to identify trends & patterns within the data.
- comparison of subsets of data, helpful in understanding performance variations across categories, regions etc.
- helps to eliminate irrelevant data & improves data accuracy.
- data visualisation like graphs, charts or tables can be customized to display specific subsets of data.
- helpful for gaining insights.

Sorting Reports / Data sorting.

- process of arranging / organizing set of data in a particular order / sequence.
- involves rearranging the data elements based on certain criteria to make it easier to analyze, search or present the information.
- can be done in ascending or descending order, depending on the requirements or nature of data.
- criteria for sorting depends on the data type & purpose of analysis.
- common sorting criteria include alphabetical order, numerical value, data & time etc.
- allows easier identification of trends & patterns.
- brings together similar elements & helps in conducting comparative analysis.
- sorting data makes it easier to search.
- sorted data is easier to present & understand.
- allows clearer visualisation of trends, ranking etc.
- help to identify & resolve inconsistencies / errors in dataset
- systematic way for checking missing values, duplicates & other quality issues.

Grouping select row/column
1) click Subtotal → Data tab.
2) Info on spreadsheet is automatically selected & subtotal dialog box opens
3) After applying subtotal on Total cost cost of col., subtotals are

1) select cell in column to be sorted
2) click sort & filter in Editing in Home tab.
3) Select A to Z. Info in category column is sorted in alpha. order
4) Similarly we can sort data in descending order.

Adding Calculations to Reports.

- reports are a combination of categorical & numerical data.
- to get summarized view of available data, we add calculations or formulas.
- helps to understand data in a better way.
- following are the calculations we can perform:
  - SUM → select cells/range → sum of all nos. in column
  - AVG → select cells/col/range → avg. of all nos. in column
  - COUNT → select cells/range → count of items in columns
  - MAX → select cells/range → highest numeric val in column
  - MIN → ——→ ——→ → lowest numeric/alpha. value
- can be used to analyze trends across different categories
- can include calculations to forcast future trends.
- can be employed to normalize data.
- In Excel sheets:
  - Autosum is used to quickly sum a column or row.
  1) Select a cell next to the nos. you want to sum.
  2) Click Autosum on home tab.
  3) When we click autosum it automatically enters formula to sum the numbers.

OR.

We can write formula instead of using Autosum function.

1) Select cell where we want calc. results & paste formula.
3) After entering we will get result of calculation.

Conditional Formatting.

- helps analyst to extract interesting data from reports.
- works on changing the appearance of cells by highlighting them in different colors or format.
- used to change the appearance of cell.
- conditions are user specified rules like comparing with some numerical values, result of some formula & text matching.

1) - We define the conditions/rules which determine when formatting should be applied.
2) We select the cells / range of cells we want to apply the conditional formatting job to.
3) Choose formatting options that should be applied to cells to meet the specified conditions
changing font color, background color, borders, font style etc.
4) Before finalizing, preview to see how it will look on selected cells.

- can be used to identify trends & outliers..
- enables visualization of data patterns / relationships.

Adding summary lines to reports.

- provides concise overview of key information.
- de condense detailed detail into summarized values.
- allows readers to grasp the main insights without delving into the entire report.
- provides a quick snapshot of most crucial info, saving reader's time.
- useful when reports are lengthy or contain complex datasets.
- some softwares automate calculation & placement of

summary lines.

1) Identify key metrics of information.
can be totals, averages, percentages, counts etc.
2) Calculate the summary values
calc may involve aggregating data, stat analysis etc
3) Decide where summary line should appear in the report.
typically at the beginning or end of report sections, highlighting most important findings & providing a recap.
4) Apply formatting to lines to make them visually distinct
using bold or larger font styles, diff font colors etc
5) Along with summary values provide contextual information.
6) Update summary lines as needed.

Drill - up.

- process of moving from a detailed level of data to higher summarized level.
(low → high granularity)
- involves aggregating / grouping data to provide broader perspective.
- In data visualisation tools like charts or dashboards, drill up allows users to interactively explore data at different levels of granularity.
- Eg. start with bar-chart showing sales data for individual months & by drilling up, you can switch to viewing data at yearly level.
- reverse of drilling down.
- commonly used when dealing with hierarchical structures
- provides users with the control over level at which they want to view data.
- enable users to view data in different contexts

- often used in OLAP tools or data visualization platforms.
- users can switch between detailed & summarized views of data to perform analysis at different levels of granularity.
- provides flexibility & multidimensional exploration of data.
- helps in understanding relationships & hierarchy within data.

1) Start at detailed level - ind. points/records are available.
2) Aggregate / group the detailed data based on certain category
3) As we drill up, level of abstraction increases.
4) Allows to gain broader perspective on data.
5) Explore different levels.

Drill Down.

- process of moving from summarized higher level of data to a more detailed level.
- involves breaking down aggregated / grouped data to explore specific subsets or individual data points.
- high granularity → low granularity.

1) Start at summarized level
2) Identify areas of interest.
3) Drill down to detailed data.
4) Gain a granular perspective.
5) Iterative exploration.

- commonly used in OLAP tool, BI tools etc.
- allows users to interact with data & dynamically drill down to lower levels of detail.
- enables users to perform detailed analysis & uncover patterns within the data.
- helps decision-making.

By clicking on visualisation

Drill through.

- accessing more detailed info. about summarized view.
- allows users to navigate from high-level summary to a more detailed view of data by accessing underlying resources.
- allows users to access more detailed levels such as individual records.
- implemented in data visualisation, BI platforms etc.

Drill up, down & through in Excel.

1) Create pivot table/chart in Excel
2) Specify range of categories on which operations will be implemented
3) It will show pivot table & chart.
4) As per level of heirarchy we can drill up or drill down data.
5) We can drill up multiple layers (levels) at a time
Right-click the item we want to drill up & click drill up/drill down & pick the level we want to drill up.
If the data items in the pivot table are grouped, we can drill up to a group name.

Run/Schedule Reports.

- reports represent extracted information from data as per specified criteria.
- organizes data in an easy to read format
- we can save it in many formats such as pdf, xml, excel, csv etc & print it directly as well.
- as per size & features of app we can generate any no. of reports as per requirement.

- Eg - if we are on page of students details of any institution, then for that role it can create report of student count, performance & merit.
After that if we switch to the next page like placement it generates a report for year-wise placement, company wise placement etc.
- We can schedule a report ~~after~~ when we create it.
- We have to choose the schedule action & then enter info such as save, print, data & time.
- When report is processed, item will be removed from the job queue.
- From the Job Queue entries page, we can change report parameters such as type of O/P file, runtime etc.

1) Click on search icon, enter Job Queue Entries, select the related entries.
2) On the Job Queue Entries page, select req: report
3) Choose set on Hold
4) Open & edit status by selecting status (on hold). Edit report
5) After editing, repeat first two steps & set the status to Ready to resume generating the report.

Different Output forms.

1) PDF.

- Portable Document File.
- type of document developed by Adobe.
- purpose was to introduce standard for rep. of document into a format independent of application software, hardware & OS.

- full capability to contain information like text, images, hyperlinks, digital signatures etc.
- In most cases existing docs are converted to Pdfs
- can contain links, buttons, form fields, audio, video & business logic
- they can be signed electronically.
- file format that captures all the elements of a printed document as an electronic image that users can view, navigate, print or forward.
- can be created using Adobe Acrobat.
- file contains one or more image pages, each of which users can zoom in or out from. scroll back & forward.
- useful when users want to preserve the original formatting of a document.
- useful when users want to send a document digitally.
- when user wants to created a document which can't easily editable.

Excel.

- one of the oldest & widely used spreadsheet software.
- file ext : added suffix to the file which specifies the file type
  - includes 2-4 characters & separated with a dot sign
  Example.docx.
- there are wide range of extensions.
- each extension provides necessary info. about the file type.
- each extension has a specific role
- not all extensions are supported with every software

- Excel file extensions help us to know more about file type & data it contains / carries

Excel File Extensions.

1) XLSX.
- default extension of Excel. (no macros used)
- based on XML
- primarily used for Excel 2007 & 2010.
- Before 2007 it was XSL.
- after introduction of ~~XSLXS~~ XLSX, file size was reduced as compared to XSL.
- downloading files became easier than before (sharing).
- free of malicious code.
↑ because XLSX cannot store VBA macro code.

2) CSV.
- comma separated values
- it stores plain text data by delimiting data entries with commas.
- can be opened in text editors, spreadsheet programs like Excel.
- plaintext file containing list of data.
- used to exchange data bet^n applications.
- simple structure
- data separated by commas.
- designed in a way to easily export data & import it into other programs.

XML.
- Extensible Markup Language.
- simple text based format for representing structured info. docs, data, config. etc.

- used for sharing info bet^n programs, people, computers & people across networks & locally.
- similar to HTML
- syntax is struct.
- very verbose.
- self describing.