



asma paracha 3/18/2024 ↻

Question and answer

Post your questions below

↻ ASMA PARACHA 3/18/24 2:18PM

What are the main differences between transactional and analytical data ? Provide examples for each.

🗨 7

Anonymous 3/18/24 3:51PM

Transactional data contains information on transactions, listing purchases and prices and perhaps the date and customer. Transaction databases are designed to be fast to handle multiple entries every day.

Analytical data is data from the past, aggregating lots of information to find trends.

Anonymous 3/18/24 3:56PM

Transactional data is constantly changing, where as Analytical data rarely changes.

Anonymous 3/18/24 3:57PM

Based on the differences you have mentioned how is data governance differs between these two types of data ?

Anonymous 3/18/24 4:15PM

transactional data focuses on real-time operations, while analytical data enables historical analysis and insights

Anonymous 3/18/24 4:16PM

Transaction highly concurrency and optimization, data consistency

Anonymous 3/18/24 4:17PM

Analytical secure storage, clean and complete data encryption

Anonymous 3/18/24 4:19PM

transaction: constant changing and usage, need controls in place

analytical: needs secure storage/encryption, make sure of data integrity and availability

↻ ASMA PARACHA 3/18/24 2:21PM

How can you ensure effective data governance during the processing phase?

🗨 3

Anonymous 3/18/24 4:44PM

The processing phase is the phase where the collected data is handled and transformed after collection. We can ensure effective data governance with data minimization techniques such as only processing necessary data, and omitting unused or PII data.

Anonymous 3/18/24 4:46PM

using template, more than often those with data entry position is not aware of all the rule/policy. so it better to go with template and example for them to follow

Anonymous 3/18/24 4:56PM

Data lineage
Data quality
Data classification
Meta data management

↻ ASMA PARACHA 3/18/24 2:44PM

What are the key security and privacy concerns associated with cloud data storage?"

10

Anonymous 3/18/24 3:50PM

When stored on the cloud, a third party has access to your data, and it can pose more risks.

Anonymous 3/18/24 3:52PM

But how are you going to deal with that risk?

Anonymous 3/18/24 3:53PM

The first thing you should do is to choose a reliable cloud service platform that follow data governance policies. You can also add a layer of protection to your data by encrypting it.

Anonymous 3/18/24 3:54PM

Encryption at the time of storage as well as processing time.

Anonymous 3/18/24 3:55PM

Yes

Anonymous 3/18/24 3:55PM

Cloud storage is outside your control and your are sharing resources with third parties.

Anonymous 3/18/24 3:56PM

Is it better than on-premise storage ? how ?

Anonymous 3/18/24 4:00PM

Cloud storage is better than on-premise storage because it leverage location limitation, it allows more computation power, it provides the ability to discover label and catalog items, it also simplify monitoring and visualization with dashboards.

Anonymous 3/18/24 4:21PM

I was thinking about AWS services such as CloudWatch that provides dashboard

Anonymous 3/18/24 4:24PM

Availability of different types of tools by the cloud provider such as data governance, monitoring and security tools.

ASMA PARACHA 3/18/24 2:47PM

Explain identity based control.

2

Anonymous 3/18/24 4:04PM

Identity based control is to ensure the resources are managed efficiently and to protect sensitive information. Here the access to resources are granted based on the identity of user , depending on their roles and responsibilities.

Anonymous 3/18/24 4:04PM

This is a type of security whereby credentials need to be given before accessing resources.

ASMA PARACHA 3/18/24 2:48PM

What is perimeter-based security and why is it failing?

2

Anonymous 3/18/24 4:02PM

Perimeter-based security is when onsite or on-network devices are trusted and able to access all functions. You secure the ability to get into premise. It's failing because once hacker get on to the network, they gain the ability to access everything and is able to do a lot of damage.

Anonymous 3/18/24 4:03PM

Perimeter-based security focuses on protecting the network perimeter from external threats. It's failing due to the rise of cloud computing, remote work, and sophisticated cyber attacks, making traditional perimeter defenses inadequate to protect.

ASMA PARACHA 3/18/24 2:50PM

What are the main advantages of zero trust model ?

2

Anonymous 3/18/24 3:51PM

Zero trust model helps to reduce the risk of data breach. Every device is authenticated and user is verified before connecting to network.

Anonymous 3/18/24 5:16PM

Enhanced security with strict access
Improved visibility and control
Reduced risk and insider threats
Securely work from anywhere

ASMA PARACHA 3/18/24 2:56PM

How data lineage help in data governance
?What are the different types of lineages?

3

Anonymous 3/18/24 3:57PM

Data lineage helps to track the changes happening in data at each stages of transformation. It helps for regulatory and auditing and reduces risks of data loss.

Anonymous 3/18/24 4:12PM

Data lineage helps you keep track of what happened to the data over its lifecycle. It can help quality monitoring. It can also help check who has accessed the data. The different types of lineage are They are 1. **Descriptive data lineage**, 2. **Automated data lineage**, 3. **Design lineage**, 4. **Business lineage**, 5. **Operational lineage as per my research on google.**

Anonymous 3/18/24 4:35PM

table level, column level

ASMA PARACHA 3/18/24 3:04PM

Guidelines for creating a data
life cycle management plan.

2

Anonymous 3/18/24 4:42PM

1. Know the type of data you have to deal with. 2. Understand the different lifecycle stages 3. Define policies for each stage 4. Assign data ownership 5. Define level of data sensitivity

Anonymous 3/18/24 5:01PM

Identify the data to be captured or collected
Define how the data will be organized
Document a data storage and preservation strategy
Define data policies
Define roles and responsibilities

ASMA PARACHA 3/18/24 3:06PM

What strategies can be employed to incorporate
security throughout the entire data lifecycle?

2

Anonymous 3/18/24 5:12PM

data encryption
access management

Anonymous 3/18/24 5:13PM

Data creation: obtain consent for data collection
Data processing: data encryption
Data usage: identity based control
Data storage: data encryption
Data archiving: establish archiving policies

Data purging: define clear data retention schedule based on legal requirements and needs

ASMA PARACHA 3/18/24 3:07PM

Different types of monitoring

3

Anonymous 3/18/24 3:57PM

Data Quality Monitoring, Usage and Access monitoring, Compliance Monitoring, Performance monitoring

Anonymous 3/18/24 3:58PM

Can someone elaborate on each one of the above listed monitorings?

Anonymous 3/18/24 4:47PM

Data quality monitoring is checking the quality over the lifecycle of the data. If the quality decreased, where did it start decreasing? Data usage and access monitoring helps you keep track of who has accessed the data and whether they should have that privilege or not. Compliance monitoring is about checking whether the data follows the relevant compliance. I think performance monitoring comes down to data quality monitoring

ASMA PARACHA 3/18/24 3:07PM

Data governance process

1

Anonymous 3/18/24 5:08PM

Define business objective
identify datas in the process
lay out data governance framework
implement data control
monitoring/maintaining

ASMA PARACHA 3/18/24 3:10PM

How to maintain and measure data quality ?

5

Anonymous 3/18/24 4:01PM

Audit actions on data.
Perform cleaning and preprocessing, checking for and handling missing values and outliers.
Have standards that are upheld.

Anonymous 3/18/24 4:11PM

How can we measure the quality ?

Anonymous 3/18/24 4:12PM

Techniques for maintaining ?

Anonymous 3/18/24 4:31PM

Accuracy
Completeness
Timeliness
Easily affected by outliers
Source trustworthiness

Anonymous 3/18/24 4:31PM

Scorecard
Prioritization
Annotation
Profiling
Deduplication
Outliers
Lineage tracking
Completeness
Merging datasets

ASMA PARACHA 3/18/24 3:12PM

Discuss the role of people in data governance.

4

Anonymous 3/18/24 4:11PM

Data owner, Data steward, Data analyst, Legal, Privacy tsar, Business analyst, external auditor, Csuit

Anonymous 3/18/24 4:38PM

Why and how are they relevant to DG ?

Anonymous 3/18/24 4:50PM

People are responsible for data governance. Data owners are those who own the data. Data stewards ensures the quality and accuracy of the data, the data analyst and business analyst manipulate and get insights from the data, legal, privacy tsar and the rest are mostly linked to the legal aspect

Anonymous 3/18/24 5:02PM

people,process, tools

ASMA PARACHA 3/18/24 3:12PM

What are some common risks in the data life cycle?

12

Anonymous 3/18/24 3:53PM

Data quality issues, security breaches, privacy violations, data loss

Anonymous 3/18/24 3:56PM

How can we mitigate these issues ?

Anonymous 3/18/24 3:58PM

Establish data quality standards, conduct regular data audits, and implement validation checks to ensure data accuracy. Encrypt data at rest, assign access controls based on roles to protect data from unauthorized access, theft and exposure.

Anonymous 3/18/24 3:59PM

Create clear data governance policies, roles and responsibilities

Anonymous 3/18/24 4:14PM

Some common risks in the data life cycle unauthorized access, data loss. It can be prevented with regular audits and backups.

Anonymous 3/18/24 4:27PM

stages in the data life cycle:

Anonymous 3/18/24 4:27PM

data creation

Anonymous 3/18/24 4:27PM

data processing

Anonymous 3/18/24 4:27PM

data storage

Anonymous 3/18/24 4:27PM

data usage

Anonymous 3/18/24 4:28PM

data archiving and data destruction

Anonymous 3/18/24 4:30PM

data life: how long you have to store the data and the proper disposal of that data is very important and a neglected issue.

ASMA PARACHA 3/18/24 4:05PM

Discuss the privacy and ethical implications of data used in AI versus ML models.

0

