Azure DevOps Assignment 3

Pratik Wani

- Notes

Day 3

* CI/CD :-
  - Concept central to software. It spans a whole feild of processess, testing methods and tooling methods and all provided facilitated by the Git code versioning process

* CI → test new code added, if it doesn't fit then immediately fixed it.

* CD → if code successfully passed the CI then immediately it made live using CD.

* Note :- for data engineer, this means faster, more reliable process, ensuring high-quility data is delivered consistently. and if any issue, it's caught and fixed swiftly.

* CI/CD pipeline :-
  · focuses on automating data operations and transformations
  · It merges developement, testing and operational workflows into a unified, autometed process ensures that data assets are consistently high quality and that data infrastructure evolve smoothly, even at scale.
  · CI/CD for data pipeline automation has become more criticle in ensuring the development velocity of different processes.

* **CI in data pipeline :-**

* **Automated Testing:-**
  Checks the integrated and quality of data transformation, ensuring that data is processed as expected and any error is spotted early.

* **version control :-**
  Data pipeline code stored in repos like Git, allowing tracking and managing changes

* **consistent environment :-**
  CI tools can run tests in environment that mirror production, ensuring that differences in configuration or dependencies don't introduce errors.

* **Data quality checks**
  Checks for null values, Data range violations data type mismatched or other custom quality rules.