BOFA : Introduction to

Apache Airflow & Astronomer

13 Oct 2025-17 Oct 2025

Training Session timings (UTC+5:30(IST))

* Start Time -> 10:00 AM
* Break -> 11:45 AM to 12:00 AM
* Lunch Break -> 01:15 PM to 02:15 PM
* Break -> 03:30 PM to 03:45 PM
* End time -> 06:00 PM ( 05:30 PM )

Reference material:

* Exercises
  + 1. Hello World" with BashOperator (Single Task)
  + 2. Sequential Workflow (Two Tasks)
  + 3. Parallel Tasks with PythonOperator
  + 4. DAG which is running 4 times a day at 1 pm, 2 pm , 6 pm and 7 PM
    - DAY 1 : <https://codeshare.io/5vXmWn>
  + 5. DAG Using Taskflow API using below code
    - DAY 2: <https://codeshare.io/adp9xg>
  + 6. Convert your DAG built in step 2 to Task flow API
  + 7. DAG with Catchup and Analyse same
  + 8. DAG with below Parameters
    - depends\_on\_past
    - wait\_for\_downstream
    - max\_active\_runs
  + 9. Astro Cloud Deployment
    - [Astronomer.io](http://astronomer.io) Trial account using personal email id
  + 10. Trigger DAG run operator
  + Day 3 code share : [Codeshare: Day3](https://docs.google.com/document/d/1X0d89XzQSlLxoy1J1vbkt9QYsO71-wXImjaOndNGH7g/edit?usp=sharing)
  + 11. Dynamic Task generation
    - Using list
    - Using Airflow Variables
    - Using API interactions
  + 12. Dynamic DAG generation
  + 13. Task Groups
    - Nested TG
    - Parallel TG
  + 14. Xcoms
    - Xcom with Airflow variables
  + 15. Git Integration and CI/CD
    - [Exercise 9\_ Exploring Git Integration with astronomer cloud .pdf](https://drive.google.com/file/d/1WDL3M40MV6_vaNgl7OgXmCR08QeBDS1F/view?usp=sharing)
  + 16. Airflow Integration with postgres
    - [Connecting Astronomer Cloud with PostgreSQL .pdf](https://drive.google.com/file/d/1PTiiMh0CDJi3NJUFRS4SIEfbycMJJ0Nc/view?usp=sharing)
  + 17 Custom Operators
    - [Codeshare: Day3, 4](https://docs.google.com/document/d/1X0d89XzQSlLxoy1J1vbkt9QYsO71-wXImjaOndNGH7g/edit?tab=t.0)
  + 18 SLA alarms
  + 19 AWS connection and S3 Read Write
    - [Exercise 8\_ Exploring usage of AWS S3 with astronomer cloud .pdf](https://drive.google.com/file/d/1rCtE7hjWnLIw8EkUgWtwqeRGnqVrsH3E/view?usp=sharing)
* Project
  + Your company ingests daily customer transaction data. A batch of raw data files (e.g., transaction\_20251017\_part001.csv, transaction\_20251017\_part002.csv) are uploaded to an S3 bucket. Once all files for a given day are successfully uploaded, a final trigger file, SUCCESS\_20251017, is placed in the bucket.
  + The data engineering team requires a pipeline that only starts processing after this SUCCESS file appears. The entire processing for a given day must be completed within 2 hours of starting. To ensure data integrity, a day's data cannot be processed if the previous day's run failed, and concurrent runs for the same logical date are forbidden.
* Reference links
  + <https://support.astronomer.io/hc/en-us>
  + <https://github.com/apache/airflow/tree/main/providers>
  + <https://registry.astronomer.io/>
  + <https://github.com/apache/airflow/issues>
* Additional Use cases
* Environmental Details
  + Docker desktop
* Basic commands & Installation Detail
* Data files
* Feedback link :
  + Consolidated Feedback
    - <https://survey.zohopublic.com/zs/94D9Ej>
    - <https://forms.gle/g6fqJuknG2Fc1bSL9>
  + Day5
    - https://forms.gle/pq7GvjgMZBDjT9416
* Contact details:
  + [Linkedin](https://www.linkedin.com/in/deeptansh-aggarwal/)
* Additional Req:
  + OnHold working demo using variables or Branch operator
  + SLA working demo
* Error file :
  + <https://codeshare.io/29YqYB>
  + [Error Document](https://docs.google.com/document/d/1ssuu35tzI-v4SVryH_WEbIkdmg2lt9RNXgj-hFNXOs8/edit?usp=sharing)
* Codeshare link
  + DAY 1 : <https://codeshare.io/5vXmWn>
  + DAY 2: <https://codeshare.io/adp9xg>
  + DAY 3, 4, 5: [Codeshare: Day3, 4, 5](https://docs.google.com/document/d/1X0d89XzQSlLxoy1J1vbkt9QYsO71-wXImjaOndNGH7g/edit?tab=t.0)