## **⚙️ PART 1: How to Install Apache Airflow on Windows (Safest Way)**

Airflow doesn’t run **natively** on Windows. So we’ll use **WSL (Windows Subsystem for Linux)** or **Docker** — two safe and official ways.

### **🔹 OPTION 1: Install Airflow on Windows using WSL (Recommended)**

#### **✅ Step-by-Step**

### **🧱 Step 1: Install WSL and Ubuntu**

In PowerShell (Admin):

wsl --install

Restart PC and Ubuntu will be installed.

### **📦 Step 2: Update and Install Python, pip, venv**

sudo apt update && sudo apt upgrade

sudo apt install python3-pip python3-venv -y

### **🏗 Step 3: Create and Activate Virtual Environment**

python3 -m venv airflow\_venv

source airflow\_venv/bin/activate

### **☁️ Step 4: Set Airflow Environment Variables**

export AIRFLOW\_HOME=~/airflow

### **📥 Step 5: Install Airflow (latest)**

pip install apache-airflow

Optional (with extras like PostgreSQL, Redis):

pip install apache-airflow[celery,postgres,redis]

### **🏗 Step 6: Initialize Airflow DB**

airflow db init

### **👤 Step 7: Create Admin User**

airflow users create \

--username admin \

--firstname Gowtham \

--lastname SB \

--role Admin \

--email admin@example.com \

--password admin123

### **🚀 Step 8: Start Webserver & Scheduler**

In two terminals:

# Terminal 1

airflow webserver --port 8080

# Terminal 2

airflow scheduler

Then open: [http://localhost:8080](http://localhost:8080/) Login: admin / admin123

## **🔁 PART 2: Schedule a Python Script in Airflow (DAG)**

### **📂 Step 1: Create a DAG File**

Create file: ~/airflow/dags/simple\_task.py

from airflow import DAG

from airflow.operators.python import PythonOperator

from datetime import datetime, timedelta

def my\_task():

with open("/home/ubuntu/airflow\_output.txt", "a") as f:

f.write(f"Task ran at: {datetime.now()}\n")

print("Task executed!")

default\_args = {

'owner': 'gowtham',

'retries': 1,

'retry\_delay': timedelta(minutes=2),

}

with DAG(

dag\_id='my\_first\_airflow\_dag',

default\_args=default\_args,

description='Simple Python print task',

start\_date=datetime(2024, 1, 1),

schedule\_interval='\*/2 \* \* \* \*', # Every 2 minutes

catchup=False,

) as dag:

task1 = PythonOperator(

task\_id='print\_time',

python\_callable=my\_task,

)

task1

### **✅ Step 2: Wait 1-2 minutes**

Then check:

* Airflow UI → DAGs → Enable my\_first\_airflow\_dag
* File: /home/ubuntu/airflow\_output.txt

You’ll see:

Task ran at: 2025-05-16 23:58:01

## **⚠️ OPTION 2: Use Airflow with Docker (If You Prefer Containers)**

You can also use the official Airflow Docker setup:  
 👉<https://airflow.apache.org/docs/apache-airflow/stable/howto/docker-compose/index.html>

But for beginners, WSL method is easier and more flexible.

## **🧠 Summary Table**

| **Step** | **What You Do** |
| --- | --- |
| 1 | Install WSL + Ubuntu |
| 2 | Setup virtualenv + Airflow |
| 3 | Init DB and create user |
| 4 | Create DAG file in ~/airflow/dags |
| 5 | View UI at localhost:8080 |
| 6 | Watch output in /home/ubuntu/airflow\_output.txt |

## **🎯 Final Tips**

* Use **crontab.guru** to generate Airflow schedule\_interval
* You can deploy more Python scripts in DAGs folder
* DAGs will auto-refresh every 30 seconds in Airflow UI

### **About the Author**

**Gowtham SB** is a **Data Engineering expert, educator,** **and content creator** with a passion for **big data technologies, as well as cloud and Gen AI** . With years of experience in the field, he has worked extensively with **cloud platforms, distributed systems, and data pipelines**, helping professionals and aspiring engineers master the art of data engineering.

Beyond his technical expertise, Gowtham is a **renowned mentor and speaker**, sharing his insights through engaging content on **YouTube and LinkedIn**. He has built one of the **largest Tamil Data Engineering communities**, guiding thousands of learners to excel in their careers.

Through his deep industry knowledge and hands-on approach, Gowtham continues to **bridge the gap between learning and real-world implementation**, empowering individuals to build **scalable, high-performance data solutions**.

𝐒𝐨𝐜𝐢𝐚𝐥𝐬

🎥𝐘𝐨𝐮𝐓𝐮𝐛𝐞 - https://www.youtube.com/@dataengineeringvideos

📸𝐈𝐧𝐬𝐭𝐚𝐠𝐫𝐚𝐦 - <https://instagram.com/dataengineeringtamil>

📸𝐈𝐧𝐬𝐭𝐚𝐠𝐫𝐚𝐦 - [https://instagram.com/](https://instagram.com/dataengineeringtamil)thedatatech.in

🤝𝐂𝐨𝐧𝐧𝐞𝐜𝐭 𝐟𝐨𝐫 𝟏:𝟏 - https://topmate.io/dataengineering/

💼𝐋𝐢𝐧𝐤𝐞𝐝𝐈𝐧 - https://www.linkedin.com/in/sbgowtham/

🌐𝐖𝐞𝐛𝐬𝐢𝐭𝐞 - https://codewithgowtham.blogspot.com

💻𝐆𝐢𝐭𝐇𝐮𝐛 - http://github.com/Gowthamdataengineer

💬𝐖𝐡𝐚𝐭𝐬 𝐀𝐩𝐩 - https://lnkd.in/g5JrHw8q

📧𝐄𝐦𝐚𝐢𝐥 - atozknowledge.com@gmail.com

📱𝐀𝐥𝐥 𝐌𝐲 𝐒𝐨𝐜𝐢𝐚𝐥𝐬 - <https://lnkd.in/gf8k3aCH>