

Operators

Operators define the atomic steps of work that make up a DAG. Airflow comes with many Operators that can perform common operations. Here are a handful of common ones:

- `PythonOperator`
- `PostgresOperator`
- `RedshiftToS3Operator`
- `S3ToRedshiftOperator`
- `BashOperator`
- `SimpleHttpOperator`
- `Sensor`

Task Dependencies

In Airflow DAGs:

- Nodes = Tasks
- Edges = Ordering and dependencies between tasks

Task dependencies can be described programmatically in Airflow using `>>` and `<<`

- `a >> b` means a comes before b
- `a << b` means a comes after b

```
hello_world_task = PythonOperator(task_id='hello_world', ...)
goodbye_world_task = PythonOperator(task_id='goodbye_world', ...)
...
# Use >> to denote that goodbye_world_task depends on hello_world_task
hello_world_task >> goodbye_world_task
```

Tasks dependencies can also be set with “set_downstream” and “set_upstream”

- `a.set_downstream(b)` means a comes before b
- `a.set_upstream(b)` means a comes after b

```
hello_world_task = PythonOperator(task_id='hello_world', ...)
goodbye_world_task = PythonOperator(task_id='goodbye_world', ...)
...
hello_world_task.set_downstream(goodbye_world_task)
```