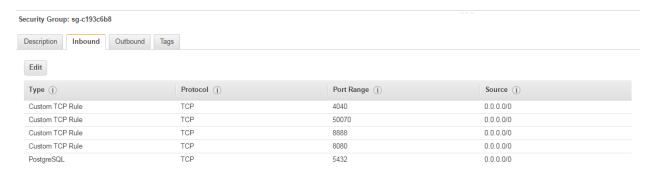
Tableau Visualization Set up Guide

Step 0: EC2 In stance configuration

Add Port (5432) to the EC2 Server security group as below:



Step 1: Starting a Hive Thrift Server for Remote Hive Access

Mount EBS volume and start Hadoop and PostgreSQL Database:

mount -t ext4 /dev/xvdf /data

/root/start-hadoop.sh

sudo -u postgres pg_ctl -D /data/pgsql_project2/data -l /data/pgsql_project2/logs/pgsql.log start

Start HiveServer2:

hive --service hiveserver2

Step 2: Connecting Tableau to PostgreSQL in EC2 Instance

Fill in information in the PostgreSQL Connection as below:

Server: EC2 Host

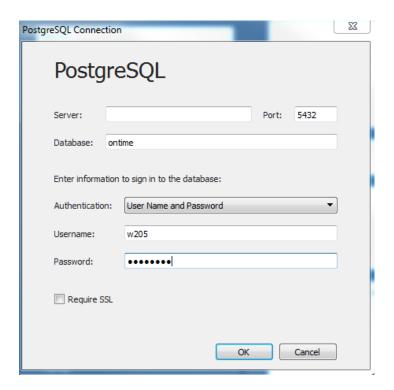
Port: 5432

Database: ontime

Authentication: Username and password

Username: w205

Password: postgres



Step 3: Loading and Joining Tables

Ontime data a

Join airplane_data b on a.tailnumber = b. tailnumber

Join airport_data c on a.Origin = c. iata

Join weather date d on a.airport = d.airport and a.FlightDate = d.Date

