Part C

1. Final Submitted Code

- a. Question 1:
 - i. PartCQuestion-1-disjoint.java
 - 1. To run: grader_dir/PartCQuestion1-1.sh
 - ii. PartCQuestion-1-joint.java
 - 1. To run: grader dir/PartCQuestion1-2.sh
- b. Question 2:
 - i. PartCQuestion-2.java
 - 1. To run: grader dir/PartCQuestion2.sh
 - 2. This application supports passing in parameter to set allowed lateness
 - 3. Eg. "flink run -c org.myorg.quickstart.late ~/software/flink-1.3.2/quickstart/target/quickstart-0.1.jar 100"
 - a. This will set allowed lateness to be 100 seconds.

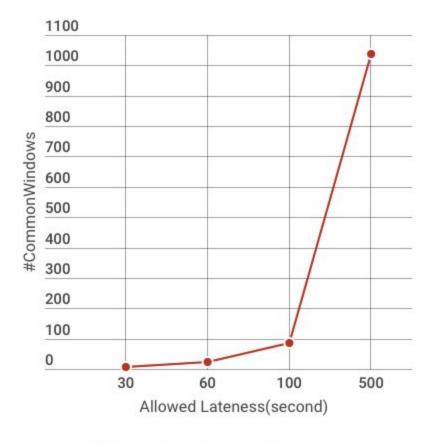
2. Output

- a. When running each application, the output will be printed to screen, as well as stored in both HDFS:/ and your working directory
- b. To note: before running single application, all previous output from this part will be deleted by default.
- c. But, all sample outputs are also stored in PartC/output/...,eg. C_Q1_disjoint.output

3. Analysis

- a. Compare the common windows between allowing lateness with allowed lateness as {30, 60, 100, 500}, and the disjoint results from Question 1
 - i. To get common windows, a script is provided: PartC/compare.sh, it mainly dose: "comm -12 <(sort \$1) <(sort \$2) > \$3"
 - ii. All common windows are stored in /PartC/output/common windows/..., eg. common 30
 - iii. Number of common windows:

	Number of Common Windows
30	5
60	23
100	85
500	1037



Number of Common Windows