2.

**Case 1:** Output of all Input files on 3 node cluster

\*\*\*\*\*The most common reason for flight cancellations considering all input files\*\*\*\*\*

Most Common Reason for Flight Cancellation is: Carrier

\*\*\*\*\*The longest and shortest average taxi time per flight (both in and out), respectively\*\*\*\*\*

The following are the 3 longest average taxi time

CKB 197.85714285714286

MKK 45.77816291161179

VLD 22.280210396885032

The following are the 3 shortest average taxi time.

BFF 1.3333333333333333

CYS 1.3333333333333333

PVU 1.7692307692307692

\*\*\*\*\*Highest and Lowest probability, respectively, for being on schedule\*\*\*\*\*

Flights with Highest probability for being on schedule

HA 0.91306132471349671

AQ 0.87046143082497837

ML (1) 0.80688378014728224

Flights with Lowest probability for being on schedule

PI 0.66221762936828543

EV 0.69313102663496834

B6 0.7088574013567475

Execution time: 16 min 34 sec Workflow Start Time: 18:12:36 Workflow End Time: 18:29:09

**Case 2:** Output of all Input files on 4 node cluster

\*\*\*\*\*The most common reason for flight cancellations considering all input files\*\*\*\*\*

Most Common Reason for Flight Cancellation is: Carrier

\*\*\*\*\*The longest and shortest average taxi time per flight (both in and out), respectively\*\*\*\*\*

The following are the 3 longest average taxi time

CKB 197.85714285714286

MKK 45.77816291161179

VLD 22.280210396885032

The following are the 3 shortest average taxi time.

BFF 1.3333333333333333

CYS 1.3333333333333333

PVU 1.7692307692307692

\*\*\*\*\*Highest and Lowest probability, respectively, for being on schedule\*\*\*\*\*

Flights with Highest probability for being on schedule

HA 0.91306132471349671

AQ 0.87046143082497837

ML (1) 0.80688378014728224

Flights with Lowest probability for being on schedule

PI 0.66221762936828543

EV 0.69313102663496834

B6 0.7088574013567475

Execution time: 13 min 35 sec Workflow Start Time: 17:40:33 Workflow End Time: 17:54:07

**Case 3:** Output of all Input files on 5 node cluster

\*\*\*\*\*The most common reason for flight cancellations considering all input files\*\*\*\*\*

Most Common Reason for Flight Cancellation is: Carrier

\*\*\*\*\*The longest and shortest average taxi time per flight (both in and out), respectively\*\*\*\*\*

The following are the 3 longest average taxi time

CKB 197.85714285714286

MKK 45.77816291161179

VLD 22.280210396885032

The following are the 3 shortest average taxi time.

BFF 1.3333333333333333

CYS 1.3333333333333333

PVU 1.7692307692307692

\*\*\*\*\*Highest and Lowest probability, respectively, for being on schedule\*\*\*\*\*

Flights with Highest probability for being on schedule

HA 0.91306132471349671

AQ 0.87046143082497837

ML (1) 0.80688378014728224

Flights with Lowest probability for being on schedule

PI 0.66221762936828543

EV 0.69313102663496834

B6 0.7088574013567475

Execution time: 12min 20sec Workflow Start Time: 17:14:51 Workflow End Time: 17:27:10

**Case 4:** Output of all Input files on 6 node cluster

\*\*\*\*\*The most common reason for flight cancellations considering all input files\*\*\*\*\*

Most Common Reason for Flight Cancellation is: Carrier

\*\*\*\*\*The longest and shortest average taxi time per flight (both in and out), respectively\*\*\*\*\*

The following are the 3 longest average taxi time

CKB 197.85714285714286

MKK 45.77816291161179

VLD 22.280210396885032

The following are the 3 shortest average taxi time.

BFF 1.3333333333333333

CYS 1.3333333333333333

PVU 1.7692307692307692

\*\*\*\*\*Highest and Lowest probability, respectively, for being on schedule\*\*\*\*\*

Flights with Highest probability for being on schedule

HA 0.91306132471349671

AQ 0.87046143082497837

ML (1) 0.80688378014728224

Flights with Lowest probability for being on schedule

PI 0.66221762936828543

EV 0.69313102663496834

B6 0.7088574013567475

Execution time: 10 min 59 sec Workflow Start Time:16:31:10 Workflow End Time:16:42:08

**Case 5:** Output of all Input files on 7 node cluster

\*\*\*\*\*The most common reason for flight cancellations considering all input files\*\*\*\*\*

Most Common Reason for Flight Cancellation is: Carrier

\*\*\*\*\*The longest and shortest average taxi time per flight (both in and out), respectively\*\*\*\*\*

The following are the 3 longest average taxi time

CKB 197.85714285714286

MKK 45.77816291161179

VLD 22.280210396885032

The following are the 3 shortest average taxi time.

BFF 1.3333333333333333

CYS 1.3333333333333333

PVU 1.7692307692307692

\*\*\*\*\*Highest and Lowest probability, respectively, for being on schedule\*\*\*\*\*

Flights with Highest probability for being on schedule

HA 0.91306132471349671

AQ 0.87046143082497837

ML (1) 0.80688378014728224

Flights with Lowest probability for being on schedule

PI 0.66221762936828543

EV 0.69313102663496834

B6 0.7088574013567475

Execution Time: 10 min 30 secs Workflow Start Time: 16:06:07 Workflow End Time:16:16:37

**Case 6:** Output of all Input files on 8 node cluster

\*\*\*\*\*The most common reason for flight cancellations considering all input files\*\*\*\*\*

Most Common Reason for Flight Cancellation is: Carrier

\*\*\*\*\*The longest and shortest average taxi time per flight (both in and out), respectively\*\*\*\*\*

The following are the 3 longest average taxi time

CKB 197.85714285714286

MKK 45.77816291161179

VLD 22.280210396885032

The following are the 3 shortest average taxi time.

BFF 1.3333333333333333

CYS 1.3333333333333333

PVU 1.7692307692307692

\*\*\*\*\*Highest and Lowest probability, respectively, for being on schedule\*\*\*\*\*

Flights with Highest probability for being on schedule

HA 0.91306132471349671

AQ 0.87046143082497837

ML (1) 0.80688378014728224

Flights with Lowest probability for being on schedule

PI 0.66221762936828543

EV 0.69313102663496834

B6 0.7088574013567475

Execution Time: 9 min 19 sec Workflow Start Time: 15:44:50 Workflow End Time: 15:54:09

3. Progressive Outputs

**Case 1:** Output of 1987.csv on 8 node cluster

\*\*\*\*\*The most common reason for flight cancellations\*\*\*\*\*

There are no cancellations according to the input files.

\*\*\*\*\*The longest and shortest average taxi time per flight (both in and out), respectively\*\*\*\*\*

There are no Taxi times for given input(s)  
There are no Taxi times for given input(s).

\*\*\*\*\*Highest and Lowest probability, respectively, for being on schedule\*\*\*\*\*

Flights with Highest probability for being on schedule

WN 0.76302490871152841

AA 0.7583959067456291

EA 0.7219411920730552

Flights with Lowest probability for being on schedule

PS 0.5738111613283553

DL 0.60019385860302005

AS 0.626649638143891

Execution time: 1 min 50 secs Workflow Start Time:12:47:21 Workflow End Time:12:49:11

**Case 2:** Output of 1987-1988.csv on 8 node cluster

\*\*\*\*\*The most common reason for flight cancellations\*\*\*\*\*

There are no cancellations according to the input files.

\*\*\*\*\*The longest and shortest average taxi time per flight (both in and out), respectively\*\*\*\*\*

There are no Taxi times for given input(s)  
There are no Taxi times for given input(s).

\*\*\*\*\*Highest and Lowest probability, respectively, for being on schedule\*\*\*\*\*

Flights with Highest probability for being on schedule

WN 0.81440074470560855

HP 0.78283514593645717

AA 0.7818557264506219

Flights with Lowest probability for being on schedule

NW 0.69221116906054

US 0.693095559678247

PI 0.6949684992693485

Execution time: 2 min 59 secs Workflow Start Time: 12:55:36 Workflow End Time: 12:58:35

**Case 3:** Output of 1987-1989.csv on 8 node cluster

\*\*\*\*\*The most common reason for flight cancellations\*\*\*\*\*

There are no cancellations according to the input files.

\*\*\*\*\*The longest and shortest average taxi time per flight (both in and out), respectively\*\*\*\*\*

There are no Taxi times for given input(s)  
There are no Taxi times for given input(s).

\*\*\*\*\*Highest and Lowest probability, respectively, for being on schedule\*\*\*\*\*

Flights with Highest probability for being on schedule  
WN 0.78548467148966103  
HP 0.77661223855994457  
AA 0.76932886495552674  
Flights with Lowest probability for being on schedule  
PI 0.66221762936828543  
US 0.66817843912617053  
UA 0.6935194135871504

Execution time: 3 min 9 secs Workflow Start Time:13:01:05 Workflow End Time: 13:04:14

**Case 4:** Output of 1987-1990.csv on 8 node cluster

\*\*\*\*\*The most common reason for flight cancellations\*\*\*\*\*

There are no cancellations according to the input files.

\*\*\*\*\*The longest and shortest average taxi time per flight (both in and out), respectively\*\*\*\*\*

There are no Taxi times for given input(s)  
There are no Taxi times for given input(s).

\*\*\*\*\*Highest and Lowest probability, respectively, for being on schedule\*\*\*\*\*

Flights with Highest probability for being on schedule  
WN 0.775724991222345  
HP 0.77530266496860376  
AA 0.76100751530067834  
Flights with Lowest probability for being on schedule  
PI 0.66221762936828543  
UA 0.704616492312905  
US 0.70722883870299846

Execution time: 3 min 10 secs Workflow Start Time: 13:06:49 Workflow End Time: 13:09:59

**Case 5:** Output of 1987-1991.csv on 8 node cluster

\*\*\*\*\*The most common reason for flight cancellations\*\*\*\*\*

There are no cancellations according to the input files.

\*\*\*\*\*The longest and shortest average taxi time per flight (both in and out), respectively\*\*\*\*\*

There are no Taxi times for given input(s)  
There are no Taxi times for given input(s).

\*\*\*\*\*Highest and Lowest probability, respectively, for being on schedule\*\*\*\*\*

Flights with Highest probability for being on schedule  
PI 0.33778237063171457  
UA 0.2865061821168094  
PS 0.2860510614511562  
Flights with Lowest probability for being on schedule  
ML (1) 0.19311621985271776  
WN 0.21770661960109144  
HP 0.2198657563381077

Execution time: 3 min 17 secs Workflow Start Time: 13:11:26 Workflow End Time: 13:14:43

**Case 6:** Output of 1987-1992.csv on 8 node cluster

\*\*\*\*\*The most common reason for flight cancellations\*\*\*\*\*

There are no cancellations according to the input files.

\*\*\*\*\*The longest and shortest average taxi time per flight (both in and out), respectively\*\*\*\*\*

There are no Taxi times for given input(s)  
There are no Taxi times for given input(s).

\*\*\*\*\*Highest and Lowest probability, respectively, for being on schedule\*\*\*\*\*

Flights with Highest probability for being on schedule  
Flights with Highest probability for being on schedule  
PI 0.33778237063171457  
PS 0.2860510614511562  
DL 0.2781941802882951  
Flights with Lowest probability for being on schedule  
WN 0.19179601339402585  
ML (1) 0.19311621985271776  
HP 0.20742706684112028

Execution time: 3 min 28 secs Workflow Start Time: 13:16:08 Workflow End Time: 13:19:36

**Case 7:** Output of 1987-1993.csv on 8 node cluster

\*\*\*\*\*The most common reason for flight cancellations\*\*\*\*\*

There are no cancellations according to the input files.

\*\*\*\*\*The longest and shortest average taxi time per flight (both in and out), respectively\*\*\*\*\*

There are no Taxi times for given input(s)  
There are no Taxi times for given input(s).

\*\*\*\*\*Highest and Lowest probability, respectively, for being on schedule\*\*\*\*\*

Flights with Highest probability for being on schedule  
PI 0.33778237063171457  
PS 0.2860510614511562  
DL 0.28279530262470254  
Flights with Lowest probability for being on schedule  
WN 0.18037741988553063  
ML (1) 0.19311621985271776  
HP 0.20640762872118035

Execution time: 3 min 31 secs Workflow Start Time: 13:22:28 Workflow End Time: 13:25:59

**Case 8:** Output of 1987-1994.csv on 8 node cluster

\*\*\*\*\*The most common reason for flight cancellations\*\*\*\*\*

There are no cancellations according to the input files.

\*\*\*\*\*The longest and shortest average taxi time per flight (both in and out), respectively\*\*\*\*\*

There are no Taxi times for given input(s)  
There are no Taxi times for given input(s).

\*\*\*\*\*Highest and Lowest probability, respectively, for being on schedule\*\*\*\*\*

Flights with Highest probability for being on schedule  
PI 0.33778237063171457  
PS 0.2860510614511562  
DL 0.2788170709538494  
Flights with Lowest probability for being on schedule  
WN 0.1808950329040245  
ML (1) 0.19311621985271776  
HP 0.2135903110932608

Execution time: 3 min 37 secs Workflow Start Time: 13:29:09 Workflow End Time: 13:32:46

**Case 9:** Output of 1987-1995.csv on 8 node cluster

\*\*\*\*\*The most common reason for flight cancellations\*\*\*\*\*

There are no cancellations according to the input files.

\*\*\*\*\*The longest and shortest average taxi time per flight (both in and out), respectively\*\*\*\*\*

The following are the 3 longest average taxi time   
JFK 16.125534471373985  
EWR 14.660862740803607  
DFW 14.176996975494419  
The following are the 3 shortest average taxi time.  
GUM 3.4761904761904763  
KSM 3.9849624060150375  
DUT 4.191292875989446

\*\*\*\*\*Highest and Lowest probability, respectively, for being on schedule\*\*\*\*\*

Flights with Highest probability for being on schedule  
WN 0.81066040295330796  
ML (1) 0.80688378014728224  
NW 0.78572804008214192  
Flights with Lowest probability for being on schedule  
PI 0.66221762936828543  
PS 0.7139489385488438  
DL 0.7195024423452635

Execution time: 3 min 52 secs Workflow Start Time: 13:35:36 Workflow End Time: 13:39:28

**Case 10:** Output of 1987-1996.csv on 8 node cluster

\*\*\*\*\*The most common reason for flight cancellations\*\*\*\*\*

There are no cancellations according to the input files.

\*\*\*\*\*The longest and shortest average taxi time per flight (both in and out), respectively\*\*\*\*\*

The following are the 3 longest average taxi time   
JFK 16.914142785124287  
EWR 15.301810349922189  
DFW 14.222427159701798  
The following are the 3 shortest average taxi time.  
GUM 3.4295302013422817  
KSM 3.802030456852792  
DUT 3.953658536585366

\*\*\*\*\*Highest and Lowest probability, respectively, for being on schedule\*\*\*\*\*

Flights with Highest probability for being on schedule  
ML (1) 0.80688378014728224  
WN 0.80358764672660388  
NW 0.7811238129773064  
Flights with Lowest probability for being on schedule  
PI 0.66221762936828543  
DL 0.7125043593430004  
PS 0.7139489385488438

Execution time: 4 min 2 secs Workflow Start Time: 13:40:31 Workflow End Time: 13:44:33

**Case 11:** Output of 1987-1997.csv on 8 node cluster

\*\*\*\*\*The most common reason for flight cancellations\*\*\*\*\*

There are no cancellations according to the input files.

\*\*\*\*\*The longest and shortest average taxi time per flight (both in and out), respectively\*\*\*\*\*

The following are the 3 longest average taxi time   
JFK 16.888491810089477  
EWR 15.672909418236278  
DFW 14.218731745521676  
The following are the 3 shortest average taxi time.  
GUM 3.4066666666666667  
KSM 3.802030456852792  
DUT 3.8598045204642637

\*\*\*\*\*Highest and Lowest probability, respectively, for being on schedule\*\*\*\*\*

Flights with Highest probability for being on schedule  
PI 0.33778237063171457  
DL 0.2874956406569996  
PS 0.2860510614511562  
Flights with Lowest probability for being on schedule  
ML (1) 0.19311621985271776  
WN 0.19641235327339612  
NW 0.2188761870226936

Execution time: 4 min 29 secs Workflow Start Time: 13:49:40 Workflow End Time: 13:54:09

**Case 12:** Output of 1987-1998.csv on 8 node cluster

\*\*\*\*\*The most common reason for flight cancellations\*\*\*\*\*

There are no cancellations according to the input files.

\*\*\*\*\*The longest and shortest average taxi time per flight (both in and out), respectively\*\*\*\*\*

The following are the 3 longest average taxi time   
JFK 17.14430383412487  
EWR 16.232759142955228  
DFW 14.06996054027045  
The following are the 3 shortest average taxi time.  
GUM 3.4066666666666667  
KSM 3.802030456852792  
DUT 3.8427768166089966

\*\*\*\*\*Highest and Lowest probability, respectively, for being on schedule\*\*\*\*\*

Flights with Highest probability for being on schedule  
PI 0.33778237063171457  
PS 0.2860510614511562  
DL 0.2859255228073407  
Flights with Lowest probability for being on schedule  
ML (1) 0.19311621985271776  
WN 0.2062188421418204  
NW 0.2299353238643991

Execution time: 5 min 6 secs Workflow Start Time: 13:55:54 Workflow End Time: 14:01:00

**Case 13:** Output of 1987-1999.csv on 8 node cluster

\*\*\*\*\*The most common reason for flight cancellations\*\*\*\*\*

There are no cancellations according to the input files.

\*\*\*\*\*The longest and shortest average taxi time per flight (both in and out), respectively\*\*\*\*\*

The following are the 3 longest average taxi time

JFK 17.348365871925527

EWR 16.649576892662143

DFW 14.05426914979551

The following are the 3 shortest average taxi time.

GUM 3.4066666666666667

DUT 3.6560636182902586

KSM 3.802030456852792

\*\*\*\*\*Highest and Lowest probability, respectively, for being on schedule\*\*\*\*\*

Flights with Highest probability for being on schedule

PI 0.33778237063171457

AS 0.2871936636560592

PS 0.2860510614511562

Flights with Lowest probability for being on schedule

ML (1) 0.19311621985271776

WN 0.21148513470870875

NW 0.22894792559530924

Execution time: 5 min 28 secs Workflow Start Time: 14:02:36 Workflow End Time: 14:08:04

**Case 14:** Output of 1987-2000.csv on 8 node cluster

\*\*\*\*\*The most common reason for flight cancellations\*\*\*\*\*

There are no cancellations according to the input files.

\*\*\*\*\*The longest and shortest average taxi time per flight (both in and out), respectively\*\*\*\*\*

The following are the 3 longest average taxi time

JFK 17.616169119728546

EWR 16.814609855571952

LGA 14.507535564452269

The following are the 3 shortest average taxi time.

GUM 3.4066666666666667

DUT 3.5764945652173914

KSM 3.802030456852792

\*\*\*\*\*Highest and Lowest probability, respectively, for being on schedule\*\*\*\*\*

Flights with Highest probability for being on schedule

PI 0.33778237063171457

AS 0.2929730557256583

PS 0.2860510614511562

Flights with Lowest probability for being on schedule

AQ 0.12013256006628004

ML (1) 0.19311621985271776

WN 0.22196567147265342

Execution time: 5 min 39 secs Workflow Start Time: 14:09:20 Workflow End Time: 14:14:59

**Case 15:** Output of 1987-2001.csv on 8 node cluster

\*\*\*\*\*The most common reason for flight cancellations\*\*\*\*\*

There are no cancellations according to the input files.

\*\*\*\*\*The longest and shortest average taxi time per flight (both in and out), respectively\*\*\*\*\*

The following are the 3 longest average taxi time

JFK 17.451648725701144

EWR 16.799488854350237

LGA 14.692579314997339

The following are the 3 shortest average taxi time.

GUM 3.4066666666666667

DUT 3.605410447761194

KSM 3.802030456852792

\*\*\*\*\*Highest and Lowest probability, respectively, for being on schedule\*\*\*\*\*

Flights with Highest probability for being on schedule

PI 0.33778237063171457

AS 0.29770491116179487

PS 0.2860510614511562

Flights with Lowest probability for being on schedule

AQ 0.17338652229662346

ML (1) 0.19311621985271776

WN 0.22053265852117898

Execution time: 6 min 7 secs Workflow Start Time: 14:16:31 Workflow End Time: 14:22:38

**Case 16:** Output of 1987-2002.csv on 8 node cluster

\*\*\*\*\*The most common reason for flight cancellations\*\*\*\*\*

There are no cancellations according to the input files.

\*\*\*\*\*The longest and shortest average taxi time per flight (both in and out), respectively\*\*\*\*\*

The following are the 3 longest average taxi time

ACY 169.5

JFK 17.35488640211598

EWR 16.670023986406115

The following are the 3 shortest average taxi time.

GUM 3.4066666666666667

DUT 3.509649820931158

KSM 3.802030456852792

\*\*\*\*\*Highest and Lowest probability, respectively, for being on schedule\*\*\*\*\*

Flights with Highest probability for being on schedule

PI 0.33778237063171457

AS 0.2936854261692632

PS 0.2860510614511562

Flights with Lowest probability for being on schedule

AQ 0.17338652229662346

ML (1) 0.19311621985271776

WN 0.2190499407450985

Execution time: 6 min 27 secs Workflow Start Time: 14:38:01 Workflow End Time: 14:44:28

**Case 17:** Output of 1987-2003.csv on 8 node cluster

\*\*\*\*\*The most common reason for flight cancellations\*\*\*\*\*

Most Common Reason for Flight Cancellation is: Carrier

\*\*\*\*\*The longest and shortest average taxi time per flight (both in and out), respectively\*\*\*\*\*

The following are the 3 longest average taxi time   
VLD 20.803848750844026  
PHF 18.289991682838924  
JFK 17.031328375157514  
The following are the 3 shortest average taxi time.  
DUT 3.366929133858268  
GUM 3.4066666666666667  
MKK 3.4390243902439024

\*\*\*\*\*Highest and Lowest probability, respectively, for being on schedule\*\*\*\*\*

Flights with Highest probability for being on schedule  
PI 0.33778237063171457  
AS 0.28721901155102914  
PS 0.2860510614511562  
Flights with Lowest probability for being on schedule  
HA 0.11242299794661191  
OO 0.13915158481329631  
AQ 0.17338652229662346

Execution time: 7 min 14 secs Workflow Start Time: 14:46:53 Workflow End Time: 14:54:07

**Case 18:** Output of 1987-2004.csv on 8 node cluster

\*\*\*\*\*The most common reason for flight cancellations\*\*\*\*\*

Most Common Reason for Flight Cancellation is: Carrier

\*\*\*\*\*The longest and shortest average taxi time per flight (both in and out), respectively\*\*\*\*\*

The following are the 3 longest average taxi time   
CKB 229.66666666666666  
MKK 45.77816291161179  
GNV 29.259280834613143  
The following are the 3 shortest average taxi time.  
PUB 1.125  
BFF 2.0  
DUT 3.357625649913345

\*\*\*\*\*Highest and Lowest probability, respectively, for being on schedule\*\*\*\*\*

Flights with Highest probability for being on schedule  
PI 0.33778237063171457  
PS 0.2860510614511562  
AS 0.28513116742196976  
Flights with Lowest probability for being on schedule  
HA 0.08113204158959901  
OO 0.16800296216123842  
AQ 0.17338652229662346

Execution time: 7 min 33 secs Workflow Start Time: 14:56:15 Workflow End Time: 15:03:48

**Case 19:** Output of 1987-2005.csv on 8 node cluster

\*\*\*\*\*The most common reason for flight cancellations\*\*\*\*\*

Most Common Reason for Flight Cancellation is: Carrier

\*\*\*\*\*The longest and shortest average taxi time per flight (both in and out), respectively\*\*\*\*\*

The following are the 3 longest average taxi time   
CKB 212.3846153846154  
MKK 45.77816291161179  
FLO 30.613598946328615  
The following are the 3 shortest average taxi time.  
PUB 0.6923076923076923  
BFF 2.0  
FMN 3.125

\*\*\*\*\*Highest and Lowest probability, respectively, for being on schedule\*\*\*\*\*

Flights with Highest probability for being on schedule  
HA 0.92665632616884555  
AQ 0.82661347770337654  
OO 0.8212630758447155  
Flights with Lowest probability for being on schedule  
PI 0.66221762936828543  
FL 0.7027403369772947  
AS 0.71130411232829815

Execution time: 7 min 58 secs Workflow Start Time: 15:04:41 Workflow End Time: 15:12:39

**Case 20:** Output of 1987-2006.csv on 8 node cluster

\*\*\*\*\*The most common reason for flight cancellations\*\*\*\*\*

Most Common Reason for Flight Cancellation is: Carrier

\*\*\*\*\*The longest and shortest average taxi time per flight (both in and out), respectively\*\*\*\*\*

The following are the 3 longest average taxi time   
CKB 197.85714285714286  
MKK 45.77816291161179  
VLD 28.221441947565545  
The following are the 3 shortest average taxi time.  
PUB 0.6  
BFF 2.0  
PVU 2.5555555555555554

\*\*\*\*\*Highest and Lowest probability, respectively, for being on schedule\*\*\*\*\*

Flights with Highest probability for being on schedule  
HA 0.92461081997533909  
AQ 0.84901531728665208  
ML (1) 0.80688378014728224

Flights with Lowest probability for being on schedule  
PI 0.66221762936828543  
EV 0.7045863741062848  
FL 0.7051060987992895

Execution time: 8 min 52 secs Workflow Start Time: 15:13:43 Workflow End Time: 15:22:35

**Case 21:** Output of 1987-2007.csv on 8 node cluster

\*\*\*\*\*The most common reason for flight cancellations\*\*\*\*\*

Most Common Reason for Flight Cancellation is: Carrier

\*\*\*\*\*The longest and shortest average taxi time per flight (both in and out), respectively\*\*\*\*\*

The following are the 3 longest average taxi time   
CKB 197.85714285714286  
MKK 45.77816291161179  
VLD 24.602860715960606  
The following are the 3 shortest average taxi time.  
PUB 0.5294117647058824  
BFF 1.3333333333333333  
CYS 1.3333333333333333

\*\*\*\*\*Highest and Lowest probability, respectively, for being on schedule\*\*\*\*\*

Flights with Highest probability for being on schedule  
HA 0.92229514238330805  
AQ 0.86740635108344058  
ML (1) 0.80688378014728224  
Flights with Lowest probability for being on schedule  
PI 0.66221762936828543  
EV 0.6883638173556771  
AS 0.70914362391097246

Execution time: 9 min 5 secs Workflow Start Time: 15:23:32 Workflow End Time: 15:32:37

**Case 22:** Output of 1987-2008.csv on 8 node cluster

\*\*\*\*\*The most common reason for flight cancellations\*\*\*\*\*

Most Common Reason for Flight Cancellation is: Carrier

\*\*\*\*\*The longest and shortest average taxi time per flight (both in and out), respectively\*\*\*\*\*

The following are the 3 longest average taxi time

CKB 197.85714285714286

MKK 45.77816291161179

VLD 22.280210396885032

The following are the 3 shortest average taxi time.

BFF 1.3333333333333333

CYS 1.3333333333333333

PVU 1.7692307692307692

\*\*\*\*\*Highest and Lowest probability, respectively, for being on schedule\*\*\*\*\*

Flights with Highest probability for being on schedule

HA 0.91306132471349671

AQ 0.87046143082497837

ML (1) 0.80688378014728224

Flights with Lowest probability for being on schedule

PI 0.66221762936828543

EV 0.69313102663496834

B6 0.7088574013567475

Execution time: 9 min 19 secs Workflow Start Time: 15:34:01 Workflow End Time: 15:43:20