

# Data for "Traffic Jam", 27-29 may 2014

## 1 Background

Description of a subset of the data used by the Qlikview applications that Skånetrafiken uses internally. The origin of the data is the realtime system Pubtrans, databases DOI and ROI and the ticketing system, database table STAT\_LOAD. These databases contains hundreds of tables so this is only a small part of the data available.

## 2 Some things to keep in mind

### 2.1 Line number

The same line number is used in several places so Skånetrafiken uses internal line numbers. To show information to the customer the internal line number is useless. Use the line name and designation from the line.csv table if useful information isn't available directly in the table.

### 2.2 Stop point

The global ID for stop point, 'StopPointGid', can be used to get some more information about a stop point from the table stoppoint.csv.

### 2.3 Date/traffic date

Normally Skånetrafiken uses traffic dates, OperatingDayDate, that starts at 04:00 in the morning. So for journeys starting after midnight and forward until 04:00 the traffic date is not the same as the normal date. Example: a journey that starts at 01:30 2014-05-08 belongs to traffic date 2014-05-07. The traffic date is the same for the whole journey.

### 2.4 Washing

In big quantities of data from complex systems there normally exists some measured values that doesn't reflect the reality. This kind of values must be removed if the data is going to be used for statistics. This is normally not a trivial task and requires a lot of system knowledge which is impossible to communicate in a short document like this one.

Some errors comes from temporary changes in the journey route for example by road work. Sometimes the vehicle makes a detour for other reasons. The equipment in the vehicle can be out of order. The stop point can have wrong geographical information entered in the system. The radio coverage can be insufficient. The vehicle might have the wrong journey information downloaded in the vehicle computer. If one searches through the data it's possible to find busses that drives back wards or over 200 km/h which of course hasn't happened in reality.

Manually a lot of these errors are easy to spot but it's harder to remove them automatically.

### 3 Format

The data is stored in comma separated value (CSV) file format. The fields are quoted (with double quotes) if needed, e.g for numbers that uses ',' as the decimal separator. Coding is UTF-8 with DOS-style line endings (CR/LF).

### 4 Arrival and departure times

The file stoptimes.csv holds information for all arrivals and departures since 2014-01-01.

#### 4.1 Short description of fields

**OperatingDayDate** Traffic date

**LineNumber** Internal line number

**DirectionOfLineGid** Global ID for direction of line

**JourneyNumber** Journey number

**StopName** Stop point name

**StopPointGid** Global ID for stop point

**LinkLength** Length to next stop point

**LastStop** 1 if last stop on journey

**TimingStop** 1 if it's a timing stop

**SequenceNumber** Sequence number for the stop point on the journey

**TimetabledEarliestDepDateTime** Earliest allowed planed time for departure from the stop point

**ArrDateTime** Arrival datetime. The time is days as a decimal number.

**DepDateTime** Departure datetime. The time is days as a decimal number.

**ArrivalState** Arrival state as a number

**DepartureState** Departure state as number

**ArrivalStateStr** Arrival state, English name

**DepartureStateStr** Departure state, English name

## 4.2 Examples

OperatingDayDate	LineNumber	DirectionOfLineGid	JourneyNumber	StopName
2014-01-01	802	9014012080210000	01145	København Østerport
2014-01-01	803	9014012080310000	01005	København Østerport
2014-01-01	803	9014012080320000	01144	Lund C
2014-01-01	130	9014012013010000	231	Malmö Södervärn
2014-01-01	130	9014012013020000	234	Lund Bankgatan
2014-01-01	802	9014012080210000	01007	København Østerport
2014-01-01	803	9014012080310000	01011	København Østerport
2014-01-01	804	9014012080420000	11002	Lund C

StopPointGid	LinkLength	LastStop	TimingStop	SequenceNumber
9022012045011001		1	1	12
9022012045011001		1	1	12
9022012081216044		1	1	12
9022012080120007		1	1	33
9022012081156002		1	1	33
9022012045011001		1	1	12
9022012045011001		1	1	12
9022012081216046		1	1	5

TimetabledEarliestDepDateTime	ArrDateTime	DepDateTime	ArrivalState
"0,17986111110804"			2
"0,20763888888905"			2
"0,20694444444234"	"0,20902777777519"	"0,21009054398019"	6
"0,20138888889051"			2
"0,20486111110949"	"0,20913194444438"		6
"0,22152777777956"			2
"0,24930555555329"			2
"0,22083333333285"	"0,22013888888614"	"0,22120193287265"	6

DepartureState	ArrivalStateStr	DepartureStateStr
2	EXPECTED	EXPECTED
2	EXPECTED	EXPECTED
9	ARRIVED	DEPARTED
2	EXPECTED	EXPECTED
6	ARRIVED	ATSTOP
2	EXPECTED	EXPECTED
2	EXPECTED	EXPECTED
9	ARRIVED	DEPARTED

## 4.3 Arrival- and departure states

The states that can be found in the fields, ArrivalState, DepartureState, ArrivalStateStr and DepartureStateStr.

### Arrivalstates

NOTEXPECTED	0
NOTCALLED	1
EXPECTED	2
CANCELLED	3
INHIBITED	4
DUE	5
ARRIVED	6
PASSED	10
MISSED	11
REPLACED	12
ASSUMEDARRIVED	13

#### Departurestates

NOTEXPECTED	0
NOTCALLED	1
EXPECTED	2
CANCELLED	3
INHIBITED	4
ATSTOP	6
BOARDING	7
BOARDINGCLOSED	8
DEPARTED	9
PASSED	10
MISSED	11
REPLACED	12
ASSUMEDDEPARTED	13

## 5 Punctuality aggregated per month

### 5.1 Short field description

Monthly aggregated punctuality data since January 2010 can be found in punctuality.csv.

**Year** Year

**Month** Month: jan, feb, mar,...

**MonthNum** Month number, 1-12

**MonthName** month and year, e.g. sep 2010

**LineNumber** Internal line number, e.g. 37

**LineName** Line name, e.g. Stadsbuss Malmö

**LineDesignation** Local line designation, ex 37

**LineNameDesignation** Line name & designation, e.g. Stadsbuss Malmö 37

**Linedirection** Internal line number and direction number separated by ' / ' e.g. 37 / 2

**DirectionOfLineName** e.g. Lindängen

**Direction** Direction: 1 or 2

**DirectionOfLineGid** Global ID for direction of line

**Early** Number of early departures

**Punctual** Number of punctual departures

**Late** Number of late departures

**VeryLate** Number of very late departures

**StopInSkåne** Set to "Inom Skåne" if the stop point is inside the region of Skåne.

**LastStop** Set to 1 if it's the last stop on the journey

## 5.2 Examples

Year	Month	MonthNum	MonthName	LineNumber	LineName	LineDesignation
2014	mar	3	mar 2014	90	MalmöExpressen TEST	90
2014	mar	3	mar 2014	90	MalmöExpressen TEST	90
2014	mar	3	mar 2014	90	MalmöExpressen TEST	90
2014	mar	3	mar 2014	90	MalmöExpressen TEST	90
2014	jan	1	jan 2014	110	Flygbussarna	110
2014	jan	1	jan 2014	110	Flygbussarna	110
2014	feb	2	feb 2014	110	Flygbussarna	110
2014	feb	2	feb 2014	110	Flygbussarna	110

LineNameDesignation	LineDirection	DirectionOfLineName	Direction
MalmöExpressen TEST 90	90 / 1	Svågertorp	1
MalmöExpressen TEST 90	90 / 1	Svågertorp	1
MalmöExpressen TEST 90	90 / 2	Nobinas depå	2
MalmöExpressen TEST 90	90 / 2	Nobinas depå	2
Flygbussarna 110	110 / 1	Malmö Airport	1
Flygbussarna 110	110 / 1	Malmö Airport	1
Flygbussarna 110	110 / 1	Malmö Airport	1
Flygbussarna 110	110 / 1	Malmö Airport	1

DirectionOfLineGid	Early	Punctual	Late	VeryLate	StopInSkåne	LastStop
9014012009010000	0	0	0	0	Inom Skåne	0
9014012009010000	0	0	0	0	Inom Skåne	1
9014012009020000	0	0	0	0	Inom Skåne	0
9014012009020000	0	0	0	0	Inom Skåne	1
9014012011010000	0	0	0	0	Inom Skåne	0
9014012011010000	0	0	0	0	Inom Skåne	1
9014012011010000	0	0	0	0	Inom Skåne	0
9014012011010000	0	0	0	0	Inom Skåne	1

## 6 Boarded - data from the ticketsystem

Ticket information data since 2014-01-01 can be found in tickets.csv.

Unfortunately the information is not linked to the correct stop point. A lot of the boarded are lumped together on a single point per journey. It's for that reason not possible to use the information to see how many that have boarded on a specific stop point. They can for example be lumped together on a "collection point" like Malmö stad.

Data from the ticketing system doesn't use traffic date, it uses normal date.

## 6.1 Short field description

**ConId** Contractor ID

**DirectionOfLineGid** Global ID for direction of line

**Direction** Direction

**LineNumber** Internal line number

**JourneyNumber** Journey number

**StopPointGid** Global ID for stop point, derived from PNT\_NUM

**SL\_MunicipalNumber** Municipal number (Kommunkod)

**ZoneNumber** Zone number

**PNT\_TYP** Point type: 1=Zone, 2=Stop, 3=Zone Set

**PNT\_NUM** Point number, a point can be a stop point, retailer, vending machine,... Stop points have numbers greater than 10 000 and less than 10000000

**ON\_NUM** Cash : The number of boarded that hasn't used card

**ON\_NUM\_CRD** Card : The number of boarded that has used card

**ON\_NUM\_PCK** Packet : The number of packet tickets

**ON\_NUM\_REG** Plus key : Holds the number of boarded that used the the plus key

**ON\_TRANSFER** Cash transfer

**ON\_TRANSFER\_CRD** Card transfer

**ON\_TRANSFER\_PCK** Packet transfer

**Boarded** Number of boarded

**NormalDate** Date (it's a normal date not a traffic date)

Boarded is the sum: "ON\_NUM" + "ON\_NUM\_CRD" + "ON\_NUM\_REG" + "ON\_TRANSFER" + "ON\_TRANSFER\_CRD".

## 6.2 Examples

ConId	DirectionOfLineGid	Direction	LineNumber	JourneyNumber
463120	9014012000110000	1	1	217
463120	9014012000110000	1	1	219
463120	9014012000110000	1	1	219
463120	9014012000110000	1	1	219
463120	9014012000110000	1	1	219
463120	9014012000110000	1	1	219
463120	9014012000110000	1	1	219
463120	9014012000110000	1	1	221

StopPointGid PNT_NUM	SL_MunicipalNumber	ZoneNumber	PNT_TYP
9022012080720001 8072001	80		2
9022012080100005 8010055	80		2
9022012080115001 8011551	80		2
9022012080120001 8012051	80		2
9022012080140001 8014051	80		2
9022012080680003 8068053	80		2
9022012080682001 8068251	80		2
		2050	1
2050			

ON_NUM	ON_NUM_CRD	ON_NUM_PCK	ON_NUM_REG	ON_TRANSFER
0	1	0	0	0
0	9	0	0	0
0	2	0	0	0
0	5	0	0	0
0	2	0	0	0
0	1	0	0	0
0	1	0	0	0
0	0	0	0	0

  

ON_TRANSFER_CRD	ON_TRANSFER_PCK	Boarded	NormalDate
0	0	1	2012-12-01
0	0	9	2012-12-01
0	0	2	2012-12-01
0	0	5	2012-12-01
0	0	2	2012-12-01
0	0	1	2012-12-01
0	0	1	2012-12-01
0	0	0	2012-12-01

## 7 Miscellaneous smaller tables

### 7.1 stoppoint.csv - Hållplatslägen

#### 7.1.1 Short field description

**StopPointGid** Global ID for stop point

**StopPointName** Stop point name

**StopPointNameWithLocal** Stop point name and local (läge) stop separated by '/'.

**StopPointDesignation** Designation for the local

**StopPointLocalNumber** The local as a number

**StopPointLocalNumberStr** The local as a number or string

**StopPointNumber** The number of the stoppoint

### 7.1.2 Examples

StopPointGid	StopPointName	StopPointNameWithLocal
9022012030099001	Staffanstorp Rondellen	Staffanstorp Rondellen / A
9022012085132031	Örtofta station	Örtofta station / 31
9022012085132032	Örtofta station	Örtofta station / 32
9022012083194031	Kattarps station	Kattarps station / 31
9022012083194032	Kattarps station	Kattarps station / 32
9022012083195031	Maria station	Maria station / 31
9022012083195032	Maria station	Maria station / 32
9022012083196031	Ödåkra station	Ödåkra station / 31
9022012083196032	Ödåkra station	Ödåkra station / 32

StopPointDesignation	StopPointLocalNumber	StopPointLocalNumberStr	StopPointNumber
A	1	A	30099001
1	31	31	85132031
2	32	32	85132032
1	31	31	83194031
2	32	32	83194032
1	31	31	83195031
2	32	32	83195032
1	31	31	83196031
2	32	32	83196032

## 7.2 line.csv - Linjer

### 7.2.1 Short field description

**LineGid** Global ID for line

**LineNumber** Internal line number

**LineName** Line name

**LineDesignation** Line designation

**LineDefaultTransportModeCode** Transport mode code. Values BUS, TRAM, TRAIN, TAXI and FERRY.

### 7.2.2 Examples



LineGid	LineNumber	LineName	LineDesignation	LineDefaultTransportModeCode
9011012000100000	1	Stadsbuss Malmö	1	BUS
9011012000200000	2	Stadsbuss Malmö	2	BUS
9011012000300000	3	Stadsbuss Malmö	3	BUS
9011012000400000	4	Stadsbuss Malmö	4	BUS
9011012000500000	5	Stadsbuss Malmö	5	BUS
9011012000600000	6	Stadsbuss Malmö	6	BUS
9011012000700000	7	Stadsbuss Malmö	7	BUS
9011012000800000	8	Stadsbuss Malmö	8	BUS

### 7.3 linedirection.csv - Linje/riktningar

#### 7.3.1 Short field description

**DirectionOfLineGid** Global ID for direction of line

**DirectionCode** Direction, possible values 1 or 2

**DirectionOfLineName** Name of the direction of line

**LineGid** The Global ID for the line

**Line\_Direction** Internal line number and direction separated by '/'

#### 7.3.2 Examples

DirectionOfLineGid	DirectionCode	DirectionOfLineName	LineGid	Line_Direction
9014012000110000	1	Kristineberg	9011012000100000	1 / 1
9014012000120000	2	Elinelund	9011012000100000	1 / 2
9014012000210000	1	Västra hamnen	9011012000200000	2 / 1
9014012000220000	2	Lindängen	9011012000200000	2 / 2
9014012000310000	1	Ringlinjen (moturs)	9011012000300000	3 / 1
9014012000320000	2	Ringlinjen (medurs)	9011012000300000	3 / 2
9014012000410000	1	Centralen	9011012000400000	4 / 1
9014012000420000	2	Klagshamn	9011012000400000	4 / 2

### 7.4 pntnum.csv - PNT\_NUM to name

#### 7.4.1 Short field description

**DATSETID** An ID not useful during traffic jam

**PNTNAM** Point name

**PNTNAMPRN** point name printed

**PNTNUM** Point number. The same as PNT\_NUM in ticket.csv.

**PNTTYP** Point type: 1=Zone, 2=Stop, 3=Zone Set. The same as PNT\_TYP in ticket.csv.

#### 7.4.2 Examples

DATSETID	PNTNAM	PNTNAMPRN	PNTNUM	PNTTYP
2794	"MÖLLE"	NULL	1001	1
2794	"NYHAMNSL/VÄSBY"	NULL	1002	1
2794	"HÖGANÄS"	NULL	1003	1
2791	"LUND C."	NULL	10	2
2791	"HJÄRUP."	NULL	11	2
2791	"ÅKARP."	NULL	12	2
2794	"1.zon 3221/3171"	NULL	1	3
2794	"22."	NULL	22	3