

Historic Bus Occupancy Data Documentation

Contents

1	Introduction	3
1.1	What is Opal?	3
2	Data overview	3
3	Using the data.....	3
3.1	COLUMN/FIELD DATA DESCRIPTIONS	3
3.2	DATA SOURCES	4
3.2.1	PTIPS - Public Transport Information and Prioritisation System....	4
3.2.2	Opal	4
3.2.3	3. Bus Fleet Capacity	4
3.3	DATA ISSUES.....	4
3.3.1	Actual Travel Time	5
3.3.2	Opal Readers.....	5
3.3.3	Bus Capacity.....	5

Author: Transport for NSW
Date: September 2017
Version: 3

1 Introduction

1.1 What is Opal?

Opal is the smartcard ticketing system which is used on Transport for NSW's public transport. Opal cards are smartcard tickets that are used to pay for travel on public transport. The Opal network covers train, bus, ferry and light rail services in Sydney, the Blue Mountains, Central Coast, Hunter, Illawarra and Southern Highlands.

2 Data overview

There are 3 weeks of data provided in this dataset.

- Monday 21 November 2016 – Sunday 27 November 2016 (post fare reform)
- Monday 26 December 2016 – Sunday 1 January 2017 (school holidays and NYE)
- Monday 8 August to Sunday 14th August 2016 (post MST retirement)

3 Using the data

3.1 COLUMN/FIELD DATA DESCRIPTIONS

CALENDAR_DATE - date on which the scheduled service ran.

ROUTE - a collection of services that are grouped together.

ROUTE_VARIANT - each route variant represents a unique sequence of stops and points travelled by a service operating on a particular route.

DIRECTION - direction of bus along a route - Inbound, Outbound, Loop.

TIMETABLE_HOUR_BAND - scheduled hour band within which the trip is either scheduled to arrive or depart.

TRIP_CODE – code showing the VEHICLETRIP number and time of arrival or departure. Inbound trips are visualised at the time the service arrives at the final (destination) transit stop. Outbound & loop trips are from the time the service departs from the first transit stop.

VEHICLETRIP - a scheduled running of a route variant eg: the 8:15 Bondi to Burwood Monday to Friday service. The same number is often applied to the same trip each day.

TRIP_POINT - denotes the start (TRIP ORIGIN) and end (TRIP DESTINATION) of the trip.

TRANSIT_STOP - the identification number for the noted transit stop/bus stop. Usually two or three digits preceded by the post code where the stop is located.

TRANSIT_STOP_SEQUENCE - sequence of the Trip in the shift - starting at 1 at the TRIP ORIGIN and ascending in numerical order to the TRIP DESTINATION.

TRANSIT_STOP_DESCRIPTION – TRANSIT_STOP + text description of transit stop location (street/ road, suburb)

TIMETABLE_TIME - the time the bus is scheduled to depart the transit stop.

ACTUAL_TIME - the time the bus actually departed the transit stop, if available.

SUBURB - suburb in which the transit stop exists.

LATITUDE - geographic locator to map noted transit stop.

LONGITUDE - geographic locator to map noted transit stop.
DEPOT - location of the depot from which the VEHICLETRIP is run.
BUS_CONFIGURATION - configuration of the bus.
BUS_DOORS - amount of doors on the bus.
SEATED_CAPACITY - seated capacity of the bus/vehicle.
STANDING_CAPACITY - standing capacity of the bus/vehicle.
TOTAL_CAPACITY - total capacity (sum of standing and seated capacity) of the bus/vehicle.
OPAL_RECORD_STATUS - indicates whether Opal records have been collected for the trip.
TIME_PERIOD – indicates sample week:
Monday 8 August to Sunday 14th August 2016
Monday 21 November 2016 – Sunday 27 November 2016
Monday 26 December 2016 – Sunday 1 January 2017

CAPACITY_BUCKET – indicates occupancy status at each bus stop per trip:

If occupancy on the bus is MORE THAN the seating capacity of the bus ----->

“STANDING_ROOM_ONLY”

If occupancy on the bus is LESS THAN 50% of the seating capacity of the bus ----->

“MANY_SEATS_AVAILABLE”

Else -----> “FEW_SEATS_AVAILABLE”

3.2 DATA SOURCES

Data provided in this extract is from three main sources:

3.2.1 PTIPS - Public Transport Information and Prioritisation System

PTIPS utilises GPS technology and timetable data to allow real-time service monitoring, identifying buses as running early or late. This travel time data is presented over the course of a trip(s).

3.2.2 Opal

Tap On and Tap Off records are matched to the transit stops (bus stop) over the course of a route/trip. Capacity Buckets (occupancy status) are also drawn from these counts.

3.2.3 3. Bus Fleet Capacity

Bus fleet data is linked to Opal data to show the seated, standing and total capacity of a trip(s). Tap On and Tap Off records are aggregated to these capacity counts.

3.3 DATA ISSUES

While data between the above-mentioned datasets is matched as comprehensively as possible some issues are present, these include:

3.3.1 Actual Travel Time

Data outages in on-board GPS systems can mean that actual travel time data for part or all of some trips maybe missing.

3.3.2 Opal Readers

Occasionally Opal readers on buses are switched off or malfunction. Where this occurs Tap On and/ or Tap Off records are not collected.

3.3.3 Bus Capacity

Where bus records do not match the Bus Fleet Capacity data table a default capacity of 45 seated and 15 standing (the most prevalent capacity format) is applied.