

Competition Task

Competition Scenario:

Part 1

1. Transform data from multiple data sources, ensure the accuracy and integrity of relevant data sets.
2. Design, build and test of a data warehouse solution.

Part 2

3. Access, analyse and interpret data to deliver easy to consume and actionable insights for the business with a dashboard report.

Part 3

Build PowerBI dashboard, convert the existing data to meaningful information with your insight.

My Analysis:

- A. I have standardized the data
- B. Created database in SSMS and transferred them
- C. Created the load database
- D. Created Fact and dim tables in to data warehouse and transferred the data tool used SSAS

Standardized the data:

SchoolCod	SchoolName	Street	Suburb	Postcod	StudentN	Indiger	LboteP	ICSEAV	LevelOfSch	Selective
1001	Abbotsford Public School	350 Great North Rd	Abbotsford	2046	602		38	1127	Primary Schoo	Not Selec
1042	Annandale North Public School	206 Johnston St	Annandale	2038	398		16	1174	Primary Schoo	Not Selec
1061	Artarmon Public School	McMillan Rd	Artarmon	2064	1125		75	1186	Primary Schoo	Not Selec
1065	Ashfield Public School	Liverpool Rd	Ashfield	2131	370		75	1066	Primary Schoo	Not Selec
1104	Seaforth Public School	37 Kempbridge Avenue	Seaforth	2092	566		18	1152	Primary Schoo	Not Selec
1166	Beecroft Public School	98 Beecroft Rd	Beecroft	2119	817		54	1199	Primary Schoo	Not Selec
1179	Bellevue Hill Public School	Victoria Rd	Bellevue Hill	2023	611		61	1153	Primary Schoo	Not Selec
1185	Belmore North Public School	Burwood Rd	Belmore	2192	340		88	964	Primary Schoo	Not Selec
1205	Berowra Public School	Hillcrest Rd	Berowra	2081	395		12	1095	Primary Schoo	Not Selec
1215	Beverly Hills Public School	Stoney Creek & King Geor	Beverly Hills	2209	527		83	1050	Primary Schoo	Not Selec
1292	Bondi Beach Public School	Campbell Pde	Bondi Beach	2026	604		25	1144	Primary Schoo	Not Selec
1320	Boronia Park Public School	113 Pittwater Rd	Hunters Hill	2110	488		25	1131	Primary Schoo	Not Selec
1323	Botany Public School	1076 Botany Rd	Botany	2019	266		48	1059	Primary Schoo	Not Selec
1358	Brighton-Le-Sands Public School	35 Crawford Road	Brighton-Le-Sands	2216	753		79	1059	Primary Schoo	Not Selec
1372	Bronte Public School	Hewlett St	Bronte	2024	430		17	1162	Primary Schoo	Not Selec
1379	Brookvale Public School	Old Pittwater Rd	Brookvale	2100	246		54	1037	Primary Schoo	Not Selec

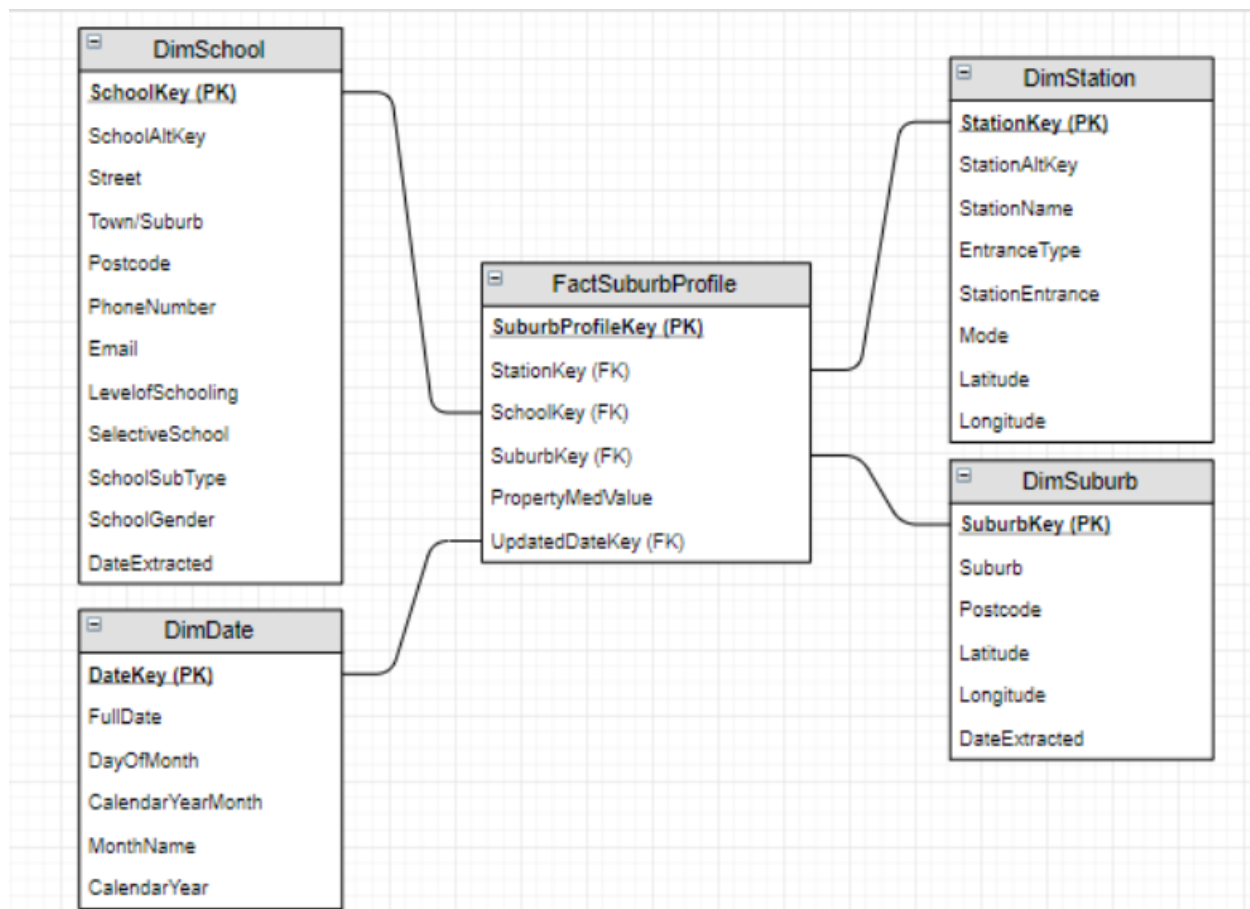
Station Entrance Dataset:

StationID	StationName	StationEntrance	EntranceType	Latitude	Longitude	Mode
1	Aberdeen	Aberdeen Macquene St Stairs	Stairs	-32.16689969	150.8919751	Train
2	Aberdeen	Aberdeen Macquene St Ramp	Ramp	-32.16688641	150.8919573	Train
3	Adamstown	Adamstown Park Ave Path	Path	-32.9337056	151.7204518	Train
4	Adamstown	Adamstown St James Rd Stairs	Stairs	-32.93341396	151.7203632	Train
5	Adamstown	Adamstown St James Rd Ramp	Ramp	-32.93340162	151.7203474	Train
6	Adamstown	Adamstown Park Ave Path	Path	-32.93382729	151.7202364	Train
7	Albion Park	Albion Park Burroo St Ramp	Ramp	-34.56340933	150.79939	Train
8	Albion Park	Albion Park Princes Hwy Ramp	Ramp	-34.56350712	150.7991893	Train
9	Albion Park	Albion Park Princes Hwy Stairs	Stairs	-34.56293995	150.798574	Train
10	Albion Park	Albion Park Princes Hwy Ramp	Ramp	-34.56278202	150.7984185	Train
11	Allawah	Allawah Railway Pde Ramp	Ramp	-33.96996874	151.1149448	Train
12	Allawah	Allawah Roberts Lane Ramp	Ramp	-33.96953727	151.1146945	Train
13	Allawah	Allawah Roberts Lane Stairs	Stairs	-33.96949011	151.1145667	Train
14	Allawah	Allawah Railway Pde Lift	Lift	-33.96987782	151.1144352	Train
15	Allawah	Allawah Railway Pde Ramp	Ramp	-33.96988689	151.1142405	Train

Property Median Value Dataset:

StateName	StateCode	City/Town	Suburb	Postcode	PropertyMedianValue	UpdatedDate
NEW SOUTH WALES	NSW	SYDNEY	Dawes Point	2000	\$1,044,350.00	1/06/2017
NEW SOUTH WALES	NSW	SYDNEY	Haymarket	2000	\$1,044,350.00	1/06/2017
NEW SOUTH WALES	NSW	SYDNEY	Millers Point	2000	\$1,044,350.00	1/06/2017
NEW SOUTH WALES	NSW	SYDNEY	Parliament House	2000	\$1,044,350.00	1/06/2017
NEW SOUTH WALES	NSW	SYDNEY	Sydney	2000	\$1,044,350.00	1/06/2017
NEW SOUTH WALES	NSW	SYDNEY	Sydney South	2000	\$1,044,350.00	1/06/2017
NEW SOUTH WALES	NSW	SYDNEY	The Rocks	2000	\$1,044,350.00	1/06/2017
NEW SOUTH WALES	NSW	SYDNEY	Broadway	2007	\$1,051,540.00	1/06/2017
NEW SOUTH WALES	NSW	SYDNEY	Ultimo	2007	\$1,051,540.00	1/06/2017
NEW SOUTH WALES	NSW	SYDNEY	Chippendale	2008	\$1,052,500.00	1/06/2017
NEW SOUTH WALES	NSW	SYDNEY	Darlington	2008	\$1,052,500.00	1/06/2017
NEW SOUTH WALES	NSW	SYDNEY	Pymont	2009	\$985,000.00	1/06/2017
NEW SOUTH WALES	NSW	SYDNEY	Darlinghurst	2010	\$1,250,000.00	1/06/2017
NEW SOUTH WALES	NSW	SYDNEY	Surry Hills	2010	\$1,250,000.00	1/06/2017
NEW SOUTH WALES	NSW	SYDNEY	Elizabeth Bay	2011	\$850,000.00	1/06/2017

Data warehouse design



- [-] 📁 Project Key DW
 - [+] 📁 Database Diagrams
 - [-] 📁 Tables
 - [+] 📁 System Tables
 - [+] 📁 FileTables
 - [+] 📁 External Tables
 - [+] 📁 Graph Tables
 - [+] 📄 **dbo.Dim Date**
 - [+] 📄 dbo.DimGeography
 - [+] 📄 dbo.DimSchool
 - [+] 📄 dbo.Dimstation
 - [+] 📄 dbo.Fact_PropertymedianValue

```

USE [Project Key DW]
GO

/***** Object: Table [dbo].[Dim Date]    Script Date: 4/07/2022 6:38:13 PM *****/
SET ANSI_NULLS ON
GO

SET QUOTED_IDENTIFIER ON
GO

CREATE TABLE [dbo].[Dim Date](
    [DateKey] [int] IDENTITY(1,1) NOT NULL,
    [FullDate] [datetime] NULL,
    [DateName] [nvarchar](255) NULL,
    [DayOfWeek] [float] NULL,
    [DayNameOfWeek] [nvarchar](255) NULL,
    [DayOfMonth] [float] NULL,
    [DayOfYear] [float] NULL,
    [WeekdayWeekend] [nvarchar](255) NULL,
    [WeekOfYear] [float] NULL,
    [MonthName] [nvarchar](255) NULL,
    [MonthOfYear] [float] NULL,
    [IsLastDayOfMonth] [nvarchar](255) NULL,
    [CalendarQuarter] [float] NULL,
    [CalendarYear] [float] NULL,
    [CalendarYearMonth] [nvarchar](255) NULL,
    [CalendarYearQtr] [nvarchar](255) NULL,
    [FiscalMonthOfYear] [float] NULL,
    [FiscalQuarter] [float] NULL,
    [FiscalYear] [float] NULL,
    [FiscalYearMonth] [nvarchar](255) NULL,
    [FiscalYearQtr] [nvarchar](255) NULL,
    CONSTRAINT [PK_Dim Date] PRIMARY KEY CLUSTERED
(
    [DateKey] ASC
)WITH (PAD_INDEX = OFF, STATISTICS_NORECOMPUTE = OFF, IGNORE_DUP_KEY = OFF, ALLOW_ROW_LOCKS = ON, ALLOW_PAGE_LOCKS = ON, OPTIMIZE_FOR_SEQUENTIAL_KEY = OFF) ON [PRIMARY]
) ON [PRIMARY]
GO

```

```

USE [Project Key DW]
GO

/***** Object: Table [dbo].[DimGeography]    Script Date: 4/07/2022 6:40:44 PM *****/
SET ANSI_NULLS ON
GO

SET QUOTED_IDENTIFIER ON
GO

CREATE TABLE [dbo].[DimGeography](
    [geokey] [int] IDENTITY(1,1) NOT NULL,
    [geoalternatekey] [int] NULL,
    [suburb] [nvarchar](255) NULL,
    [postcode] [int] NULL,
    [city] [nvarchar](255) NULL,
    [lat] [float] NULL,
    [lon] [float] NULL,
    [state_code] [nvarchar](255) NULL,
    CONSTRAINT [PK_DimGeography] PRIMARY KEY CLUSTERED
(
    [geokey] ASC
)WITH (PAD_INDEX = OFF, STATISTICS_NORECOMPUTE = OFF, IGNORE_DUP_KEY = OFF, ALLOW_ROW_LOCKS = ON, ALLOW_PAGE_LOCKS = ON, OPTIMIZE_FOR_SEQUENTIAL_KEY = OFF) ON [PRIMARY]
) ON [PRIMARY]
GO

```

```

[school_name] [nvarchar](255) NULL,
[street] [nvarchar](255) NULL,
[postcode] [float] NULL,
[student_number] [float] NULL,
[indigenous_pct] [float] NULL,
[lbote_pct] [float] NULL,
[ICSEA_Value] [float] NULL,
[level_of_schooling] [nvarchar](255) NULL,
[selective_school] [nvarchar](255) NULL,
[opportunity_class] [nvarchar](255) NULL,
[school_specialty_type] [nvarchar](255) NULL,
[school_subtype] [nvarchar](255) NULL,
[support_classes] [nvarchar](255) NULL,
[preschool_ind] [nvarchar](255) NULL,
[distance_education] [nvarchar](255) NULL,
[intensive_english_centre] [nvarchar](255) NULL,
[school_gender] [nvarchar](255) NULL,
[phone] [nvarchar](255) NULL,
[school_email] [nvarchar](255) NULL,
[fax] [nvarchar](255) NULL,
[late_opening_school] [nvarchar](255) NULL,
[date_1st_teacher] [nvarchar](255) NULL,
[lga] [nvarchar](255) NULL,
[electorate] [nvarchar](255) NULL,
[fed_electorate] [nvarchar](255) NULL,
[operational_directorate] [nvarchar](255) NULL,
[principal_network] [nvarchar](255) NULL,
[fac_district] [nvarchar](255) NULL,
[local_health_district] [nvarchar](255) NULL,
[aecg_region] [nvarchar](255) NULL,
[ASGS_remoteness] [nvarchar](255) NULL,
[latitude] [float] NULL,
[longitude] [float] NULL,
[date_extracted] [datetime] NULL,
CONSTRAINT [PK_DimSchool] PRIMARY KEY CLUSTERED
(
    [Schoolkey] ASC
)WITH (PAD_INDEX = OFF, STATISTICS_NORECOMPUTE = OFF, IGNORE_DUP_KEY = OFF, ALLOW_ROW_LOCKS = ON, ALLOW_PAGE_LOCKS = ON, OPTIMIZE_FOR_SEQUENTIAL_KEY = OFF) ON [PRIMARY]
) ON [PRIMARY]

```

```

USE [Project Key DW]
GO

/***** Object: Table [dbo].[Dimstation]    Script Date: 4/07/2022 6:42:13 PM *****/
SET ANSI_NULLS ON
GO

SET QUOTED_IDENTIFIER ON
GO

CREATE TABLE [dbo].[Dimstation](
    [stkey] [int] IDENTITY(1,1) NOT NULL,
    [stalternetkey] [int] NULL,
    [Train_Station] [nvarchar](255) NULL,
    [TS_Name] [nvarchar](255) NULL,
    [Entrance_Type] [nvarchar](255) NULL,
    [LAT] [float] NULL,
    [LONG] [float] NULL,
    [Exit_Number] [nvarchar](255) NULL,
    CONSTRAINT [PK_Dimstation] PRIMARY KEY CLUSTERED
(
    [stkey] ASC
)WITH (PAD_INDEX = OFF, STATISTICS_NORECOMPUTE = OFF, IGNORE_DUP_KEY = OFF, ALLOW_ROW_LOCKS = ON, ALLOW_PAGE_LOCKS = ON, OPTIMIZE_FOR_SEQUENTIAL_KEY = OFF) ON [PRIMARY]
) ON [PRIMARY]
GO

```

```

:CREATE TABLE [dbo].[Fact_PropertyMedianValue](
    [FactPropMedValueKey] [int] IDENTITY(1,1) NOT NULL,
    [geoKey] [int] NULL,
    [StKey] [int] NULL,
    [SchoolKey] [int] NULL,
    [DateKey] [int] NULL,
    [Suburb] [nvarchar](255) NULL,
    [Postcode] [float] NULL,
    [State] [nvarchar](255) NULL,
    [City/Town] [nvarchar](255) NULL,
    [District] [nvarchar](255) NULL,
    [Location] [nvarchar](255) NULL,
    [Property_Median_Value] [money] NULL,
    [Updated_Year] [float] NULL,
    [Updated_Month] [float] NULL,
    [date_extracted] [datetime] NULL,
    CONSTRAINT [PK_Fact_PropertyMedianValue] PRIMARY KEY CLUSTERED
(
    [FactPropMedValueKey] ASC
)WITH (PAD_INDEX = OFF, STATISTICS_NORECOMPUTE = OFF, IGNORE_DUP_KEY = OFF, ALLOW_ROW_LOCKS = ON, ALLOW_PAGE_LOCKS = ON, OPTIMIZE_FOR_SEQUENTIAL_KEY = OFF) ON [PRIMARY]
) ON [PRIMARY]
GO

:ALTER TABLE [dbo].[Fact_PropertyMedianValue] WITH CHECK ADD CONSTRAINT [FK_Fact_PropertyMedianValue_DimGeography] FOREIGN KEY([geoKey])
REFERENCES [dbo].[DimGeography] ([geokey])
GO

ALTER TABLE [dbo].[Fact_PropertyMedianValue] CHECK CONSTRAINT [FK_Fact_PropertyMedianValue_DimGeography]
GO

```

Power Bi report

