Technical Documentation

**Transformations & Lineage**

**[DRAFT]**

**Client:** Melbourne Business School

**Version:**v0.2

**Ref:**#0003

**File:** Test1.pbix

**URL**: https://app.powerbi.com/view?r=eyJrIjoiN2NjMjZiNTYtMDhjNi00MjMwLWIxNjEtYjkwYmZjMzYzYTNkIiwidCI6ImU3ZTAzMWZjLWY1MGEtNDA2OS05NWE5LTZmNGQ4OTgxYzdmMiJ9





**Clean & Wrangle datasets**



**‘course build data.xlsx**

|  |  |  |
| --- | --- | --- |
| Transformation ID | Applied to column(s) | Description |
| 1 | Topic Name | Replaced value “Toipc” with “Topic” |
| 2 | Topic\_Tag\_Name | Added new column called “Topic\_Tag\_Name”, filled down so that each topic was tagged with appropriate course name |
| 3 | Topic Name | Filtered course name, leaving only topics |
| 4 | In\_Development; Released; Planning | Created new columns to serve as indicators of a courses of a courses progress, e.g. if ‘Planning’ = 1 then the topic in the given row is in planning. |
| 5 | Count\_being\_worked\_on, | Created a column to flag all records which are flagged 1 in columns In\_Development; Released; Planning |
| 6 | Count\_all | Created a column which counts 1 for each record irrespective of whether it’s being worked on or not. |
| 7 | URL\_Image | Created a column containing URL’s pointed to dummy images |
| 8 | SME Name | Created a column containing dummy SME names |

let

Source = Excel.Workbook(File.Contents("**FILEPATH\course build data.xlsx**"), null, true),

Sheet1\_Sheet = Source{[Item="Sheet1",Kind="Sheet"]}[Data],

#"Changed Type" = Table.TransformColumnTypes(Sheet1\_Sheet,{{"Column1", type text}, {"Column2", type text}}),

#"Promoted Headers" = Table.PromoteHeaders(#"Changed Type", [PromoteAllScalars=true]),

#"Changed Type1" = Table.TransformColumnTypes(#"Promoted Headers",{{"Topic Name", type text}, {"Build Stage", type text}}),

#"Replaced Value" = Table.ReplaceValue(#"Changed Type1","Toipc","Topic",Replacer.ReplaceText,{"Topic Name"}),

#"Added Conditional Column" = Table.AddColumn(#"Replaced Value", "Topic\_Name\_Tag", each if not Text.Contains([Topic Name], "Topic") then [Topic Name] else null),

#"Filled Down" = Table.FillDown(#"Added Conditional Column",{"Topic\_Name\_Tag"}),

#"Filtered Rows" = Table.SelectRows(#"Filled Down", each Text.Contains([Topic Name], "Topic")),

#"Grouped Rows" = Table.Group(#"Filtered Rows", {"Topic\_Name\_Tag"}, {{"Count", each \_, type table [Topic Name=nullable text, Build Stage=nullable text, Topic\_Name\_Tag=text]}}),

#"Added Custom" = Table.AddColumn(#"Grouped Rows", "Custom.1", each Table.AddIndexColumn([Count], "Index", 1)),

#"Expanded Custom.1" = Table.ExpandTableColumn(#"Added Custom", "Custom.1", {"Topic Name", "Build Stage", "Index"}, {"Custom.1.Topic Name", "Custom.1.Build Stage", "Custom.1.Index"}),

#"Removed Columns" = Table.RemoveColumns(#"Expanded Custom.1",{"Count"}),

#"Added Conditional Column1" = Table.AddColumn(#"Removed Columns", "In\_Development", each if [Custom.1.Build Stage] = "In Development" then 1 else null),

#"Added Conditional Column2" = Table.AddColumn(#"Added Conditional Column1", "Released", each if [Custom.1.Build Stage] = "Released" then 1 else null),

#"Added Conditional Column3" = Table.AddColumn(#"Added Conditional Column2", "Planning", each if [Custom.1.Build Stage] = "Planning" then 1 else null),

#"Added Conditional Column4" = Table.AddColumn(#"Added Conditional Column3", "Count\_being\_worked\_on", each if [Custom.1.Build Stage] <> null then 1 else null),

#"Added Custom1" = Table.AddColumn(#"Added Conditional Column4", "Count\_All", each 1),

#"Changed Type2" = Table.TransformColumnTypes(#"Added Custom1",{{"Count\_being\_worked\_on", Int64.Type}, {"Count\_All", Int64.Type}}),

#"Added Conditional Column5" = Table.AddColumn(#"Changed Type2", "URL\_Image", each if [Topic\_Name\_Tag] = "STEM" then "https://st.depositphotos.com/2218212/2938/i/950/depositphotos\_29387653-stock-photo-facebook-profile.jpg" else if [Topic\_Name\_Tag] = "Digital, Data and Cyber Literacy" then "https://st.depositphotos.com/2218212/2938/i/950/depositphotos\_29388097-stock-photo-facebook-profile.jpg" else if [Topic\_Name\_Tag] = "Service Orientation" then "https://st.depositphotos.com/2218212/2938/i/950/depositphotos\_29387653-stock-photo-facebook-profile.jpg" else if [Topic\_Name\_Tag] = "Critical Thinking" then "https://st.depositphotos.com/2218212/2938/i/950/depositphotos\_29387653-stock-photo-facebook-profile.jpg" else "https://st.depositphotos.com/2218212/2938/i/950/depositphotos\_29387653-stock-photo-facebook-profile.jpg"),

#"Added Conditional Column6" = Table.AddColumn(#"Added Conditional Column5", "SME Name", each if [Topic\_Name\_Tag] = "Critical Thinking" then "Professor Thinking" else if [Topic\_Name\_Tag] = "Digital, Data and Cyber Literacy" then "Professor Data" else if [Topic\_Name\_Tag] = "Human Skills" then "Professor Human" else if [Topic\_Name\_Tag] = "Parliamentary and Media Acumen" then "Professor Parliamentary" else if [Topic\_Name\_Tag] = "Partnerships" then "Professor Partnerships" else if [Topic\_Name\_Tag] = "Policy Making" then "Professor Policy" else if [Topic\_Name\_Tag] = "Service Orientation" then "Professor Service" else if [Topic\_Name\_Tag] = "STEM" then "Professor STEM" else null)

in

#"Added Conditional Column6"

**‘dummy enrolment data.xlsx’**

|  |  |  |
| --- | --- | --- |
| Transformation ID | Applied to column(s) | Description |
| 1 | Created function Query1 | Function which replaces column name based on its position (index) in a table |
| 2 | Build\_on\_you | Change name of column using Query1 function to ‘RepeatableHeading’ |
| 3 | AgeIndex | Index based on first digit in age\_group, used to sort visualisations numerically. |

let

Source = Excel.Workbook(File.Contents(“**FILEPATH\ \dummy enrolment data.xlsx"),** null, true),

Sheet1\_Sheet = Source{[Item="Sheet1",Kind="Sheet"]}[Data],

#"Promoted Headers" = Table.PromoteHeaders(Sheet1\_Sheet, [PromoteAllScalars=true]),

#"Changed Type" = Table.TransformColumnTypes(#"Promoted Headers",{{"Curio ID", Int64.Type}, {"Canvas User ID", Int64.Type}, {"First Name", type text}, {"Last Name", type text}, {"Gender", type text}, {"working\_location", type any}, {"defence\_group", type text}, {"aps\_level", type text}, {"age\_group", type text}, {"build\_on\_you", type text}, {"accept\_feedback", type text}, {"Mobile phone", type any}, {"Email", type text}, {"Course ID", Int64.Type}, {"Course Start", type datetime}, {"Course Name", type text}, {"Enrolment date", type date}}),

#"Rename Column" = Query1(#"Changed Type", {9}, {"RepeatableHeading"}),

#"Added Custom" = Table.AddColumn(#"Rename Column", "AgeIndex", each Text.Start([age\_group], 1))

in

#"Added Custom"

**‘coach booking.csv’**

‘coaching book Q1-Q3’

|  |  |  |
| --- | --- | --- |
| Transformation ID | Applied to column(s) | Description |
| 1 | Column1 | Remove all ***other*** columns |
| 2 | Column1 | Remove top 1 row |
| 3 | Custom | Add new column called “Custom”, which repeats Column1 where string contains “ – Please” else null. Filled down so that each record in Column1 was tagged with correct question. |
| 4 | Column1 | Remove records where string contained “ – Please” |
| 5 | Custom | Filtered out Q4 records |
| 6 | Whole table | Pivot table show Q1,Q2,Q3 as columns populated by their respective answers |
| 7 | Whole table | Remove top 1 row |

let

Source = Csv.Document(File.Contents(“**FILEPATH\coach booking.csv**"),[Delimiter=",", Columns=8, Encoding=1252, QuoteStyle=QuoteStyle.None]),

#"Changed Type" = Table.TransformColumnTypes(Source,{{"Column1", type text}, {"Column2", type text}, {"Column3", type text}, {"Column4", type text}, {"Column5", type text}, {"Column6", type text}, {"Column7", type text}, {"Column8", type text}}),

#"Removed Other Columns" = Table.SelectColumns(#"Changed Type",{"Column1"}),

#"Removed Top Rows" = Table.Skip(#"Removed Other Columns",4),

#"Added Conditional Column" = Table.AddColumn(#"Removed Top Rows", "Custom", each if Text.Contains([Column1], " - Please") then [Column1] else null),

#"Filled Down" = Table.FillDown(#"Added Conditional Column",{"Custom"}),

#"Filtered Rows" = Table.SelectRows(#"Filled Down", each not Text.Contains([Column1], " - Please")),

#"Filtered Rows1" = Table.SelectRows(#"Filtered Rows", each ([Custom] <> "Q4 - Please select a coaching session") and ([Column1] <> "")),

#"Grouped Rows" = Table.Group(#"Filtered Rows1", {"Custom"}, {{"Count", each \_, type table [Column1=nullable text, Custom=text]}}),

#"Added Custom" = Table.AddColumn(#"Grouped Rows", "Custom.1", each Table.AddIndexColumn([Count], "Index", 1)),

#"Expanded Custom.1" = Table.ExpandTableColumn(#"Added Custom", "Custom.1", {"Column1", "Index"}, {"Custom.1.Column1", "Custom.1.Index"}),

#"Removed Columns" = Table.RemoveColumns(#"Expanded Custom.1",{"Count"}),

#"Pivoted Column" = Table.Pivot(#"Removed Columns", List.Distinct(#"Removed Columns"[Custom]), "Custom", "Custom.1.Column1"),

#"Removed Columns1" = Table.RemoveColumns(#"Pivoted Column",{"Custom.1.Index"}),

#"Removed Top Rows1" = Table.Skip(#"Removed Columns1",1)

in

#"Removed Top Rows1"

coaching book Q4\_1

|  |  |  |
| --- | --- | --- |
| Transformation ID | Applied to column(s) | Description |
| 1 | Column2 | Remove all rows which are blank |
| 2 | Whole table | Remove top 2 rows |
| 3 | Promote headers | Make new top row = column headings |

let

Source = Csv.Document(File.Contents("**FILEPATH\coach booking.csv"),[**Delimiter=",", Columns=8, Encoding=1252, QuoteStyle=QuoteStyle.None]),

#"Changed Type" = Table.TransformColumnTypes(Source,{{"Column1", type text}, {"Column2", type text}, {"Column3", type text}, {"Column4", type text}, {"Column5", type text}, {"Column6", type text}, {"Column7", type text}, {"Column8", type text}}),

#"Filtered Rows" = Table.SelectRows(#"Changed Type", each ([Column2] <> "")),

#"Removed Top Rows" = Table.Skip(#"Filtered Rows",2),

#"Promoted Headers" = Table.PromoteHeaders(#"Removed Top Rows", [PromoteAllScalars=true]),

#"Changed Type1" = Table.TransformColumnTypes(#"Promoted Headers",{{"#", Int64.Type}, {"Answer", type text}, {"%", Percentage.Type}, {"Count", Int64.Type}, {"", type text}, {"\_1", type text}, {"\_2", type text}, {"\_3", type text}})

in

#"Changed Type1"

coaching book Q4\_1

|  |  |  |
| --- | --- | --- |
| Transformation ID | Applied to column(s) | Description |
| 1 | Column2 | Remove all rows which are blank |
| 2 | Whole table | Remove bottom 5 rows |
| 3 | Promote headers | Make top row = column headings |

let

Source = Csv.Document(File.Contents("**FILEPATH\coach booking.csv"),[**Delimiter=",", Columns=8, Encoding=1252, QuoteStyle=QuoteStyle.None]),

#"Changed Type" = Table.TransformColumnTypes(Source,{{"Column1", type text}, {"Column2", type text}, {"Column3", type text}, {"Column4", type text}, {"Column5", type text}, {"Column6", type text}, {"Column7", type text}, {"Column8", type text}}),

#"Filtered Rows" = Table.SelectRows(#"Changed Type", each ([Column2] <> "")),

#"Removed Bottom Rows" = Table.RemoveLastN(#"Filtered Rows",5),

#"Promoted Headers" = Table.PromoteHeaders(#"Removed Bottom Rows", [PromoteAllScalars=true]),

#"Changed Type1" = Table.TransformColumnTypes(#"Promoted Headers",{{"#", Int64.Type}, {"Field", type text}, {"Minimum", Int64.Type}, {"Maximum", Int64.Type}, {"Mean", type number}, {"Std Deviation", type number}, {"Variance", type number}, {"Count", Int64.Type}})

in

#"Changed Type1"

**‘dummy zoom attendance.csv’**

No transformations required.

let

Source = Csv.Document(File.Contents("**FILEPATHStaging\_Folder\dummy zoom attendance.csv**"),[Delimiter=",", Columns=4, Encoding=65001, QuoteStyle=QuoteStyle.None]),

#"Promoted Headers" = Table.PromoteHeaders(Source, [PromoteAllScalars=true]),

#"Changed Type" = Table.TransformColumnTypes(#"Promoted Headers",{{"Name (Original Name)", type text}, {"User Email", type text}, {"Total Duration (Minutes)", Int64.Type}, {"Guest", type text}})

in

#"Changed Type"

**‘dummy zoom poll data.csv’**

|  |  |  |
| --- | --- | --- |
| Transformation ID | Applied to column(s) | Description |
| 1 | Index | Added index column |
| 2 | Custom | Where index = 3, fill with Column1 value (which contains topic name) else blank. Fill down. |
| 3 | Whole table | Remove top 6 rows |
| 4 | Column1; Index | Remove from dataset |
| 5 | Submitted Date | Create new column which shows submission by date only (as opposed to date-time) |

let

Source = Csv.Document(File.Contents("**FILEPATH\\dummy zoom poll data.csv**"),[Delimiter=",", Columns=6, Encoding=65001, QuoteStyle=QuoteStyle.None]),

#"Changed Type" = Table.TransformColumnTypes(Source,{{"Column1", type text}, {"Column2", type text}, {"Column3", type text}, {"Column4", type text}, {"Column5", type text}, {"Column6", type text}}),

#"Added Index" = Table.AddIndexColumn(#"Changed Type", "Index", 0, 1, Int64.Type),

#"Added Conditional Column" = Table.AddColumn(#"Added Index", "Custom", each if [Index] = 3 then [Column1] else null),

#"Filled Down" = Table.FillDown(#"Added Conditional Column",{"Custom"}),

#"Removed Top Rows" = Table.Skip(#"Filled Down",6),

#"Renamed Columns1" = Table.RenameColumns(#"Removed Top Rows",{{"Column2", "Username"}, {"Column3", "User Email"}, {"Column4", "Submitted Date/Time"}, {"Column5", "Question"}, {"Column6", "Answer"}, {"Custom", "Topic"}}),

#"Removed Columns" = Table.RemoveColumns(#"Renamed Columns1",{"Column1", "Index"}),

#"Changed Type1" = Table.TransformColumnTypes(#"Removed Columns",{{"Submitted Date/Time", type datetime}}),

#"Duplicated Column" = Table.DuplicateColumn(#"Changed Type1", "Submitted Date/Time", "Submitted Date/Time - Copy"),

#"Changed Type2" = Table.TransformColumnTypes(#"Duplicated Column",{{"Submitted Date/Time - Copy", type date}}),

#"Renamed Columns" = Table.RenameColumns(#"Changed Type2",{{"Submitted Date/Time - Copy", "Submitted date"}})

in

#"Renamed Columns"

**‘FSB Report.csv’**

|  |  |  |
| --- | --- | --- |
| Transformation ID | Applied to column(s) | Description |
| 1 | Column2 | Remove blank rows |
| 2 | Custom | If Column1 = 1 return column2 (identifies course name), and fill down so that each record in is tagged with correct course name. |
| 3 | Mean, Std Deviation, Variance, Count | Filled down so that the given value appears per-record rather than in only one record and blank for all others |
| 4 | Column2 | Keep only the following values: "Not at all interested", "Slightly interested", "Moderately interested", "Very Interested", "Extremely interested", "Total" |
| 5 | Column2, Column3, Column4, Column5, Column6, Column7, Column8, Custom | Rename to: Answer, Minimum (%), Maximum (count), Mean, Std Deviation, Variance, Count, Field |
| 6 | Key | Search Field column for string and return string if true, e.g. if Field contains “STEM” then return “STEM” |

let

Source = Csv.Document(File.Contents("**FILEPATH\FSB Report.csv**"),[Delimiter=",", Columns=8, Encoding=1252, QuoteStyle=QuoteStyle.Csv]),

#"Changed Type" = Table.TransformColumnTypes(Source,{{"Column1", type text}, {"Column2", type text}, {"Column3", type text}, {"Column4", type text}, {"Column5", type text}, {"Column6", type text}, {"Column7", type text}, {"Column8", type text}}),

#"Filtered Rows" = Table.SelectRows(#"Changed Type", each ([Column2] <> "")),

#"Added Conditional Column" = Table.AddColumn(#"Filtered Rows", "Custom", each if [Column1] = "1" then [Column2] else null),

#"Filled Down" = Table.FillDown(#"Added Conditional Column",{"Custom"}),

#"Reordered Columns" = Table.ReorderColumns(#"Filled Down",{"Custom", "Column1", "Column2", "Column3", "Column4", "Column5", "Column6", "Column7", "Column8"}),

#"Replaced Value" = Table.ReplaceValue(#"Reordered Columns","",null,Replacer.ReplaceValue,{"Column5"}),

#"Replaced Value1" = Table.ReplaceValue(#"Replaced Value","",null,Replacer.ReplaceValue,{"Column6"}),

#"Replaced Value2" = Table.ReplaceValue(#"Replaced Value1","",null,Replacer.ReplaceValue,{"Column7"}),

#"Replaced Value3" = Table.ReplaceValue(#"Replaced Value2","",null,Replacer.ReplaceValue,{"Column8"}),

#"Filled Down1" = Table.FillDown(#"Replaced Value3",{"Column5", "Column6", "Column7", "Column8"}),

#"Filtered Rows1" = Table.SelectRows(#"Filled Down1", each [Column2] = "Not at all interested" or [Column2] = "Slightly interested" or [Column2] = "Moderately interested" or [Column2] = "Very Interested" or [Column2] = "Extremely interested" or [Column2] = "Total"),

#"Renamed Columns" = Table.RenameColumns(#"Filtered Rows1",{{"Column3", "Minimum (%)"}, {"Column4", "Maximum (count)"}, {"Column5", "Mean"}, {"Column6", "Std Deviation"}, {"Column7", "Variance"}, {"Column8", "Count"}, {"Column2", "Answer"}, {"Custom", "Field"}}),

#"Added Conditional Column1" = Table.AddColumn(#"Renamed Columns", "Key", each if Text.Contains([Field], "Critical Thinking") then "Critical Thinking" else if Text.Contains([Field], "Digital, Data and Cyber Literacy") then "Digital, Data and Cyber Literacy" else if Text.Contains([Field], "Human Skills") then "Human Skills" else if Text.Contains([Field], "Parliamentary and Media Acumen") then "Parliamentary and Media Acumen" else if Text.Contains([Field], "Partnerships") then "Partnerships" else if Text.Contains([Field], "Policy Making") then "Policy Making" else if Text.Contains([Field], "Service Orientation") then "Service Orientation" else if Text.Contains([Field], "STEM") then "STEM" else null),

#"Changed Type1" = Table.TransformColumnTypes(#"Added Conditional Column1",{{"Minimum (%)", Percentage.Type}, {"Maximum (count)", Int64.Type}, {"Mean", Currency.Type}, {"Std Deviation", Currency.Type}, {"Variance", Currency.Type}, {"Count", Int64.Type}})

in

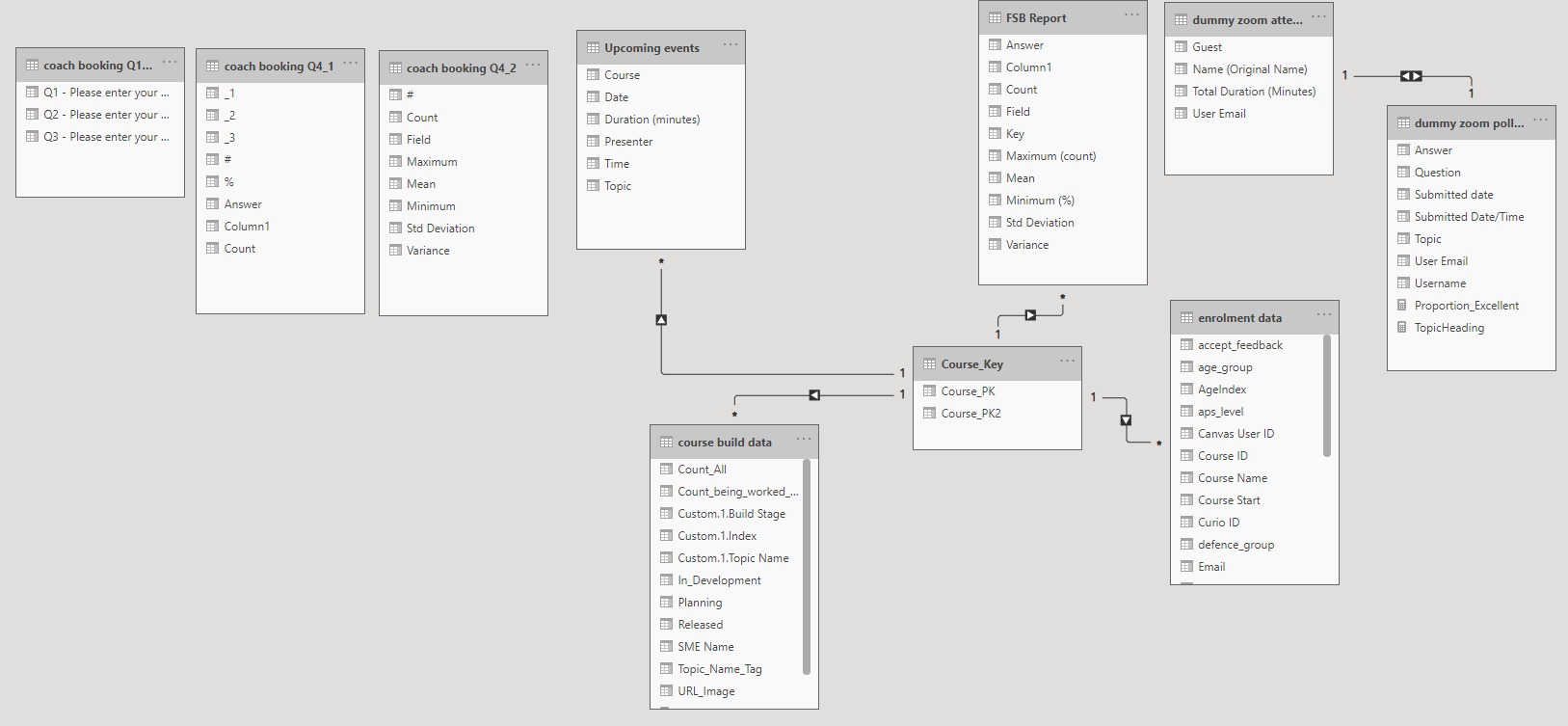
#"Changed Type1"

**Upcoming events**

Created dummy upcoming events data

**Course Key (custom PowerBI table)**

Created Course primary key to establish relationships between it and ‘course build data’, ‘enrolment data’, ‘FSB Report’, and ‘Upcoming events’ – tables are joined based on the course name.

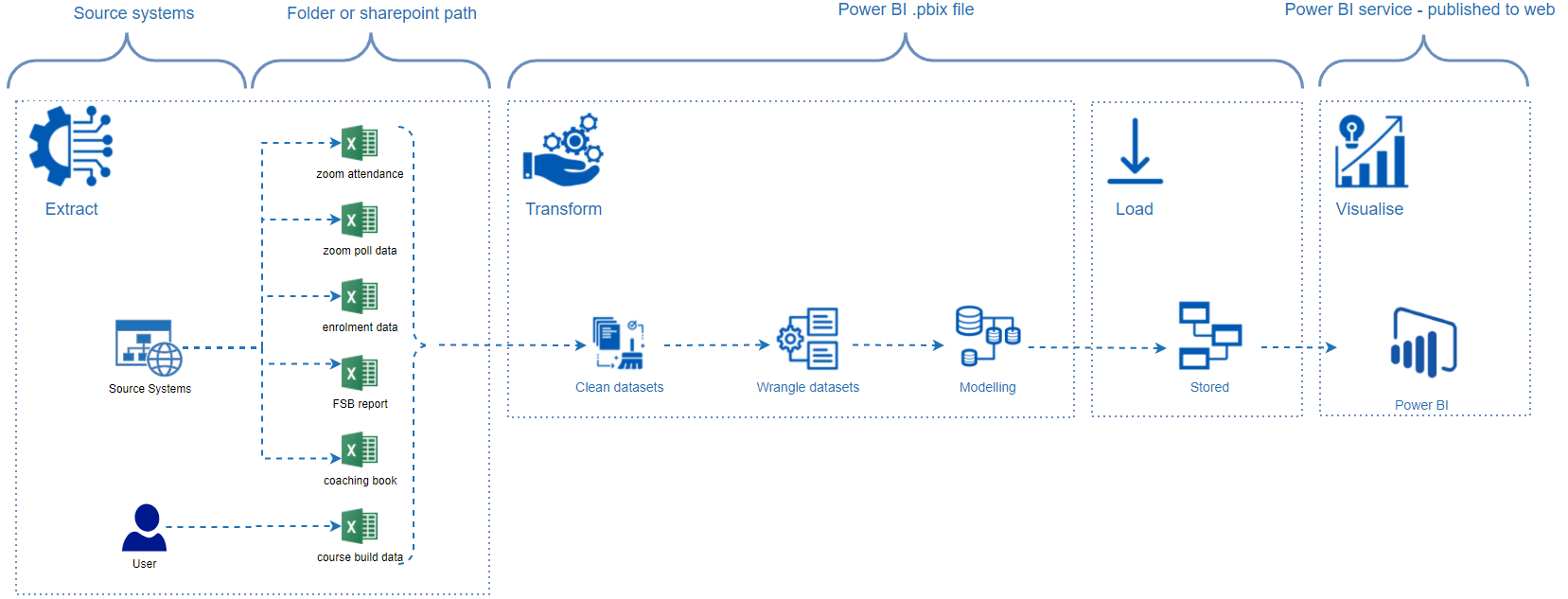
**Data modelling**

**Data modelling**



|  |  |  |
| --- | --- | --- |
| From: Table (column() | To: Table (Column) | Cardinality - Direction |
| Course build data (Topic\_Name\_Tag) | Course\_Key (**Course\_PK**) | Many to one - Single |
| Dummy zoom poll data (Username) | Dummy zoom attendance (**Name (Original Name))** | One to one - Both |
| Enrolment data (Course Name) | Course\_Key (**Course\_PK2**) | Many to one - Single |
| FSB Report (Key) | Course\_Key (**Course\_PK**) | Many to one - Single |
| Upcoming events (Course) | Course\_Key (**Course\_PK**) | Many to one - Single |

**Data Lineage**



Technical Documentation

**Wireframe**

**[DRAFT]**

**Client:** Melbourne Business School

**Version:**v0.2

**Ref:**#0003

**File:** Test1.pbix

**URL**: https://app.powerbi.com/view?r=eyJrIjoiN2NjMjZiNTYtMDhjNi00MjMwLWIxNjEtYjkwYmZjMzYzYTNkIiwidCI6ImU3ZTAzMWZjLWY1MGEtNDA2OS05NWE5LTZmNGQ4OTgxYzdmMiJ9

