

Power BI Data Transformation Documentation

This document outlines the data cleaning and transformation steps applied to the “Dirty 1” table using Power Query in Power BI. The goal was to convert an unstructured, multi-header layout into a clean, analysis-ready format.

1. Source

- Connected to the Excel file and selected the sheet named “Sheet”.

	Column1	Column2	Column3	Column4	Column5	Column6	Column7	Column8	Column9	Column10	
	Valid 100% Error 0% Empty 0%	Valid 100% Error 0% Empty 0%	Valid 9% Error 0% Empty 91%	Valid 5% Error 0% Empty 95%	Valid 2% Error 0% Empty 98%	Valid 4% Error 0% Empty 96%	Valid < 1% Error 0% Empty 99%	Valid < 1% Error 0% Empty 99%	Valid 11% Error 0% Empty 89%	Valid 6% Error 0% Empty 94%	3 0 97
1	Ship Mode	First Class		null	null	Same Day		null	null	Second Class	null
2	Segment	Consumer	Corporate	Home Office	Consumer	Corporate	Home Office	Consumer	Corporate	Home Office	
3	Order Date		null	null	null	null	null	null	null	null	
4	2013-03-14		null	null	null	null	null	null	null	null	
5	2013-12-16		null	null	null	null	null	null	129.44	null	
6	2013-06-02		null	null	null	null	null	null	null	null	
7	2013-10-21		null	null	null	null	null	null	null	null	
8	2013-08-27		null	null	null	null	null	null	13.36	null	
9	2013-11-28		null	null	null	null	null	null	null	542.34	
10	2013-03-31		null	null	null	null	null	null	null	null	
11	2013-11-21		null	null	null	null	null	null	null	null	
12	2013-11-01		null	null	null	null	null	null	null	null	
13	2013-04-05		null	null	null	null	null	null	null	4251.92	
14	2013-07-05			242.546	null	null	null	null	null	null	
15	2013-01-15	149.95		null	null	null	null	null	null	null	
16	2013-12-02	null		null	null	null	null	null	null	null	
17	2013-03-19	null		590.762	null	null	null	null	null	null	
18	2013-06-27	null		null	null	null	null	null	null	null	
19	2013-01-06	null		12.78	null	null	null	null	null	null	
20	2013-05-14	null		null	null	null	null	null	null	null	
21	2013-12-12	null		null	null	null	null	null	null	null	
22	2013-04-29	null		null	null	null	null	null	null	null	
23	2013-10-24	null		null	null	null	null	null	null	null	
24	2013-08-15	243.6		null	null	null	null	null	null	null	
25	2013-01-13	null		null	null	null	null	null	545.94	null	
26	2013-04-25	null		null	null	null	null	null	null	null	
27	2013-10-18	null		null	null	null	null	null	null	2216.8	
28	2013-01-21	null		null	null	null	null	null	25.248	null	
29	2013-08-17	null		null	null	null	null	null	null	null	
30	2013-05-05	null		47.32	null	null	null	null	null	null	
31	2013-12-16	null		null	null	null	null	null	null	null	
32											

2. Transposed Table

- Used the Transpose function to flip rows into columns.
- The original data had column headers and segment info stored vertically. Transposing made the structure suitable for tabular formatting.

	ABC 123 Column1	ABC 123 Column2	ABC 123 Column3	ABC 123 Column4	ABC 123 Column5	ABC 123 Column6	ABC 123 Column7	ABC 123 Column8	ABC 123 Column9	ABC 123 Column10
	<div>Valid 38%</div> <div>Error 0%</div> <div>Empty 62%</div>	<div>Valid 100%</div> <div>Error 0%</div> <div>Empty 0%</div>	<div>Valid 100%</div> <div>Error 0%</div> <div>Empty 0%</div>	<div>Valid 8%</div> <div>Error 0%</div> <div>Empty 92%</div>	<div>Valid 15%</div> <div>Error 0%</div> <div>Empty 85%</div>	<div>Valid 15%</div> <div>Error 0%</div> <div>Empty 85%</div>	<div>Valid 15%</div> <div>Error 0%</div> <div>Empty 85%</div>	<div>Valid 15%</div> <div>Error 0%</div> <div>Empty 85%</div>	<div>Valid 15%</div> <div>Error 0%</div> <div>Empty 85%</div>	<div>Valid 15%</div> <div>Error 0%</div> <div>Empty 85%</div>
1	Ship Mode	Segment	Order Date	2013-03-14	2013-12-16	2013-06-02	2013-10-21	2013-08-27	2013-11-28	
2	First Class	Consumer	null	null	null	null	null	null	null	
3	null	Corporate	null	null	null	null	null	null	null	
4	null	Home Office	null	null	null	null	null	null	null	
5	Same Day	Consumer	null	null	null	null	null	null	null	
6	null	Corporate	null	null	null	null	null	null	null	
7	null	Home Office	null	null	null	null	null	null	null	
8	Second Class	Consumer	null	null	129.44	null	null	13.36	null	
9	null	Corporate	null	null	null	null	null	null	542.34	
10	null	Home Office	null	null	null	null	null	null	null	
11	Standard Class	Consumer	null	null	null	605.47	null	null	null	
12	null	Corporate	null	null	null	null	788.86	null	null	
13	null	Home Office	null	91.056	null	null	null	null	null	

3. Promoted Headers

- Used “Use First Row as Headers”.

4. Changed Type

- Applied automatic data type detection to all columns.




5. Filled Down

- Applied Fill Down on the “Ship Mode” and “Segment” columns.
- These columns had merged cells in the original file, resulting in nulls after transformation. Fill Down ensures each row has complete contextual info.

	ABC 123 Ship Mode	ABC 123 Segment	ABC 123 Order Date	1.2 2013-03-14	1.2 2013-12-16	1.2 2013-06-02	1.2 2013-10-21	1.2 2013-08-27	1.2 2013-11-28	1.2 2013-11-28
	<div>Valid 100%</div> <div>Error 0%</div> <div>Empty 0%</div>	<div>Valid 100%</div> <div>Error 0%</div> <div>Empty 0%</div>	<div>Valid 0%</div> <div>Error 0%</div> <div>Empty 100%</div>	<div>Valid 8%</div> <div>Error 0%</div> <div>Empty 92%</div>	<div>Valid 8%</div> <div>Error 0%</div> <div>Empty 92%</div>	<div>Valid 8%</div> <div>Error 0%</div> <div>Empty 92%</div>	<div>Valid 8%</div> <div>Error 0%</div> <div>Empty 92%</div>	<div>Valid 8%</div> <div>Error 0%</div> <div>Empty 92%</div>	<div>Valid 8%</div> <div>Error 0%</div> <div>Empty 92%</div>	<div>Valid 8%</div> <div>Error 0%</div> <div>Empty 92%</div>
	4 distinct, 0 unique	3 distinct, 0 unique		2 distinct, 1 unique	2 distinct, 1 unique	2 distinct, 1 unique	2 distinct, 1 unique	2 distinct, 1 unique	2 distinct, 1 unique	2 distinct
1	First Class	Consumer	null	null	null	null	null	null	null	
2	First Class	Corporate	null	null	null	null	null	null	null	
3	First Class	Home Office	null	null	null	null	null	null	null	
4	Same Day	Consumer	null	null	null	null	null	null	null	
5	Same Day	Corporate	null	null	null	null	null	null	null	
6	Same Day	Home Office	null	null	null	null	null	null	null	
7	Second Class	Consumer	null	null	129.44	null	null	13.36	null	
8	Second Class	Corporate	null	null	null	null	null	null	542.34	
9	Second Class	Home Office	null	null	null	null	null	null	null	
10	Standard Class	Consumer	null	null	null	605.47	null	null	null	
11	Standard Class	Corporate	null	null	null	null	788.86	null	null	
12	Standard Class	Home Office	null	91.056	null	null	null	null	null	

6. Unpivoted Other Columns

- Selected key identifier columns and used **Unpivot Other Columns**.
- To convert wide-format columns (e.g., multiple columns for segments like Consumer, Corporate) into tidy row-based structure (attribute-value pairs).

	Ship Mode	Segment	Attribute	Value
	 4 distinct, 0 unique	 3 distinct, 0 unique	 822 distinct, 822 unique	
1	First Class	Consumer	2013-01-15	149.95
2	First Class	Consumer	2013-08-15	243.6
3	First Class	Consumer	2013-12-24	9.568
4	First Class	Consumer	2013-04-07	8.96
5	First Class	Consumer	2013-05-19	34.2
6	First Class	Consumer	2013-09-05	31.984
7	First Class	Consumer	2013-08-12	286.65
8	First Class	Consumer	2013-07-05_16	514.03
9	First Class	Consumer	2013-04-30_18	1000.95
10	First Class	Consumer	2013-03-23	9.912
11	First Class	Consumer	2013-12-30	39.128
12	First Class	Consumer	2014-04-18	106.5
13	First Class	Consumer	2014-11-21	18.176
14	First Class	Consumer	2014-12-23	194.32
15	First Class	Consumer	2014-03-23	59.48
16	First Class	Consumer	2014-10-30	182.91
17	First Class	Consumer	2014-04-16	2298.9
18	First Class	Consumer	2014-11-02	197.72
19	First Class	Consumer	2014-11-30_40	440.144
20	First Class	Consumer	2014-11-22	32.985
21	First Class	Consumer	2014-12-11	196.62
22	First Class	Consumer	2014-12-21_53	47.976
23	First Class	Consumer	2015-12-10	1056.86
24	First Class	Consumer	2015-07-25	73.2
25	First Class	Consumer	2015-02-12	1366.04
26	First Class	Consumer	2015-08-15	15.552
27	First Class	Consumer	2015-06-05	75.88
28	First Class	Consumer	2015-12-06_60	191.82
29	First Class	Consumer	2015-08-14	125.92
30	First Class	Consumer	2015-12-17	34.944
31	First Class	Consumer	2015-06-15	58.4
32	First Class	Consumer	2015-05-06_67	79.14

umn profiling based on top 1000 rows

7. Extracted Text Before Delimiter

- Extracted using text before delimiter “_” in date field.

	ABC Ship Mode	ABC Segment	ABC Attribute	ABC Value
	 4 distinct, 0 unique	 3 distinct, 0 unique	 604 distinct, 444 unique	
1	First Class	Consumer	2013-01-15	149.95
2	First Class	Consumer	2013-08-15	243.6
3	First Class	Consumer	2013-12-24	9.568
4	First Class	Consumer	2013-04-07	8.96
5	First Class	Consumer	2013-05-19	34.2
6	First Class	Consumer	2013-09-05	31.984
7	First Class	Consumer	2013-08-12	286.65
8	First Class	Consumer	2013-07-05	514.03
9	First Class	Consumer	2013-04-30	1000.95
10	First Class	Consumer	2013-03-23	9.912
11	First Class	Consumer	2013-12-30	39.128
12	First Class	Consumer	2014-04-18	106.5
13	First Class	Consumer	2014-11-21	18.176
14	First Class	Consumer	2014-12-23	194.32
15	First Class	Consumer	2014-03-23	59.48
16	First Class	Consumer	2014-10-30	182.91
17	First Class	Consumer	2014-04-16	2298.9
18	First Class	Consumer	2014-11-02	197.72
19	First Class	Consumer	2014-11-30	440.144
20	First Class	Consumer	2014-11-22	32.985
21	First Class	Consumer	2014-12-11	196.62
22	First Class	Consumer	2014-12-21	47.976
23	First Class	Consumer	2015-12-10	1056.86
24	First Class	Consumer	2015-07-25	73.2
25	First Class	Consumer	2015-02-12	1366.04
26	First Class	Consumer	2015-08-15	15.552
27	First Class	Consumer	2015-06-05	75.88
28	First Class	Consumer	2015-12-06	191.82
29	First Class	Consumer	2015-08-14	125.92
30	First Class	Consumer	2015-12-17	34.944
31	First Class	Consumer	2015-06-15	58.4
32	First Class	Consumer	2015-05-06	79.14

umn profilina based on top 1000 rows

8. Renamed Columns

- Renamed generic columns (like “Attribute”, “Value”) to meaningful names such as Date, Sales.

9. Changed Type1

- Re-applied column data types for Date as date and decimal number for Sales.

	A ^B _C Ship Mode	A ^B _C Segment	Date	1.2 Sales
	<p>Valid 100% Error 0% Empty 0%</p> <p>4 distinct, 0 unique</p>	<p>Valid 100% Error 0% Empty 0%</p> <p>3 distinct, 0 unique</p>	<p>Valid 100% Error 0% Empty 0%</p> <p>604 distinct, 444 unique</p>	<p>Valid 100% Error 0% Empty 0%</p> <p>800 distinct, 782 unique</p>
1	First Class	Consumer	2013-01-15	149.95
2	First Class	Consumer	2013-08-15	243.6
3	First Class	Consumer	2013-12-24	9.568
4	First Class	Consumer	2013-04-07	8.96
5	First Class	Consumer	2013-05-19	34.2
6	First Class	Consumer	2013-09-05	31.984
7	First Class	Consumer	2013-08-12	286.65
8	First Class	Consumer	2013-07-05	514.03
9	First Class	Consumer	2013-04-30	1000.95
10	First Class	Consumer	2013-03-23	9.912
11	First Class	Consumer	2013-12-30	39.128
12	First Class	Consumer	2014-04-18	106.5
13	First Class	Consumer	2014-11-21	18.176
14	First Class	Consumer	2014-12-23	194.32
15	First Class	Consumer	2014-03-23	59.48
16	First Class	Consumer	2014-10-30	182.91
17	First Class	Consumer	2014-04-16	2298.9
18	First Class	Consumer	2014-11-02	197.72
19	First Class	Consumer	2014-11-30	440.144
20	First Class	Consumer	2014-11-22	32.985
21	First Class	Consumer	2014-12-11	196.62
22	First Class	Consumer	2014-12-21	47.976
23	First Class	Consumer	2015-12-10	1056.86
24	First Class	Consumer	2015-07-25	73.2
25	First Class	Consumer	2015-02-12	1366.04
26	First Class	Consumer	2015-08-15	15.552
27	First Class	Consumer	2015-06-05	75.88
28	First Class	Consumer	2015-12-06	191.82
29	First Class	Consumer	2015-08-14	125.92
30	First Class	Consumer	2015-12-17	34.944
31	First Class	Consumer	2015-06-15	58.4
32	First Class	Consumer	2015-05-06	79.14

umn profiling based on top 1000 rows