# Power Query Report

1. Data Loading & Initial Setup

- Loaded CSV file from your local directory with 21 columns
- Promoted first row to column headers
- Set appropriate data types for all columns

2. Gender Standardization

- Consolidated gender categories by replacing "Non-Binary" with "Female" ("Female" Is the mode of the column)
- Replaced "Other" gender entries with "Female"
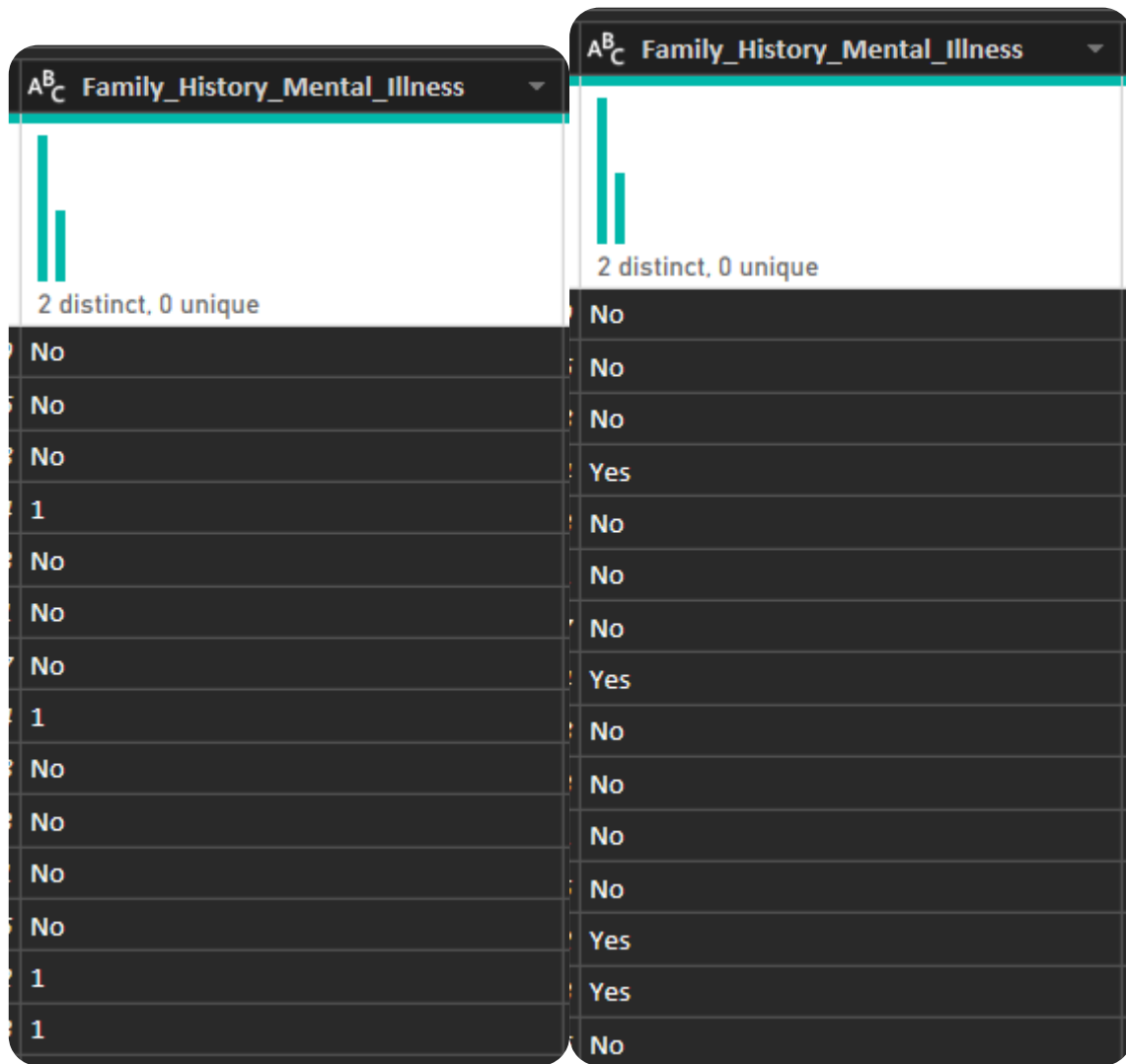- Simplified from 4 gender categories to 2 (Male/Female)
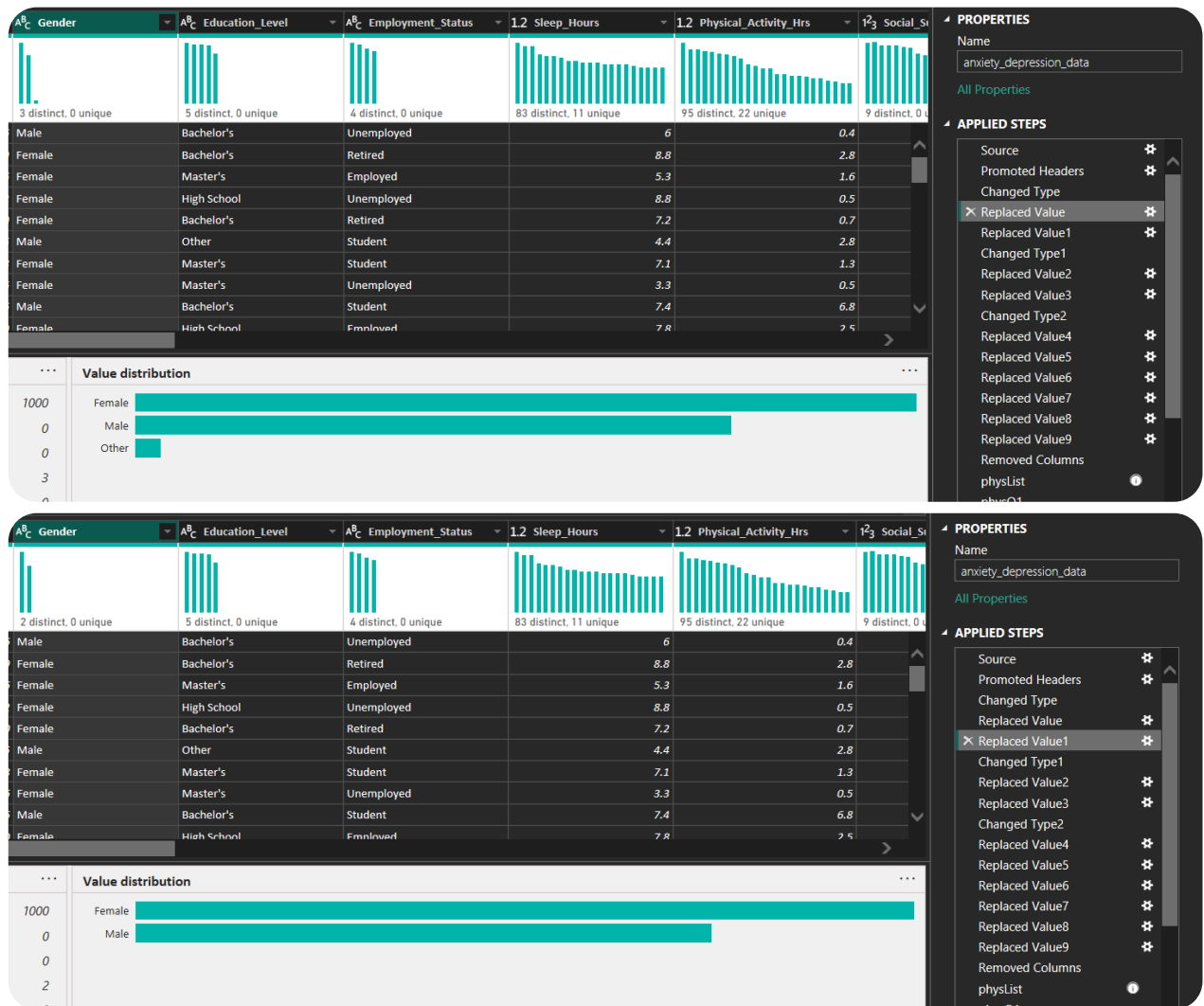
3. Binary Column Conversion

- Converted Family_History_Mental_Illness from numeric (0/1) to text (No/Yes)
- Converted Chronic_Illnesses from numeric (0/1) to text (No/Yes)
- Converted Therapy from numeric (0/1) to text (No/Yes)

- Converted Meditation from numeric (0/1) to text (No/Yes)

1²3 Family_History_Mental_Illness

2 distinct, 0 unique

0
0
0
1
0
0
0
1
0
0
0
0
1
1
0

ᴬᴮC Family_History_Mental_Illness

2 distinct, 0 unique

0
0
0
1
0
0
0
1
0
0
0
0
1
1
0
0

Family_History_Mental_Illness (left column)

2 distinct, 0 unique

Family_History_Mental_Illness
No
No
No
1
No
No
No
1
No
No
No
No
1
1

Family_History_Mental_Illness (right column)

2 distinct, 0 unique

Family_History_Mental_Illness
No
No
No
Yes
No
No
No
Yes
No
No
No
No
Yes
Yes
No

4. Column Removal

- Removed the "Substance_Use" and "Medication_Use" column (likely due to data quality issues or irrelevance)

Medication_Use	Therapy	Meditation	Substance_Use
3 distinct, 0 unique	2 distinct, 0 unique	2 distinct, 0 unique	3 distinct, 0 unique
None	No	Yes	None
None	Yes	No	None
None	No	Yes	None
None	No	No	None
None	Yes	Yes	Frequent
Occasional	No	No	Occasional
None	Yes	Yes	None
None	No	No	None
Occasional	No	No	Occasional
None	No	No	None
Occasional	No	No	Occasional
Occasional	No	No	Frequent
Occasional	Yes	No	None
None	No	No	None
None	No	Yes	Occasional
Regular	No	No	Frequent
Occasional	No	Yes	None
None	No	No	None
None	No	Yes	None
None	No	Yes	None
None	No	No	Occasional
None	No	Yes	Frequent
None	Yes	No	None
None	No	Yes	Occasional

	Family_History_Mental_Illness	Chronic_Illnesses	Therapy	Meditation
	2 distinct, 0 unique	2 distinct, 0 unique	2 distinct, 0 unique	2 distinct, 0 unique
9	No	No	No	Yes
6	No	No	Yes	No
8	No	No	No	Yes
4	Yes	Yes	No	No
3	No	No	Yes	Yes
1	No	No	No	No
7	No	No	Yes	Yes
4	Yes	Yes	No	No
8	No	Yes	No	No
3	No	No	No	No
1	No	No	No	No
5	No	No	No	No
2	Yes	No	Yes	No
3	Yes	No	No	No
6	No	No	No	Yes
3	No	No	No	No
3	No	No	No	Yes
6	No	No	No	No
5	No	Yes	No	Yes
9	No	No	No	Yes
1	No	No	No	No
3	Yes	Yes	No	Yes
3	No	No	Yes	No
6	Yes	No	No	Yes

5. Outlier Detection & Treatment

- **Physical Activity**: Calculated IQR (Q1, Q3, IQR, bounds) and replaced outliers with median
- **Sleep Hours**: Applied same IQR method and replaced outliers with median
- Created new cleaned columns for both variables

```
physList = List.RemoveNulls(Table.Column(#"Removed
Columns","Physical_Activity_Hrs")),
physQ1 = List.Percentile(physList, 0.25),
physQ3 = List.Percentile(physList, 0.75),
physIQR = physQ3 - physQ1,
physLower = physQ1 - 1.5 * physIQR,
physUpper = physQ3 + 1.5 * physIQR,
physMedian = List.Median(physList),
```

6. Data Type Optimization

- Ensured cleaned outlier columns had proper numeric data types

```
sleepList = List.RemoveNulls(Table.Column(#"Removed
```
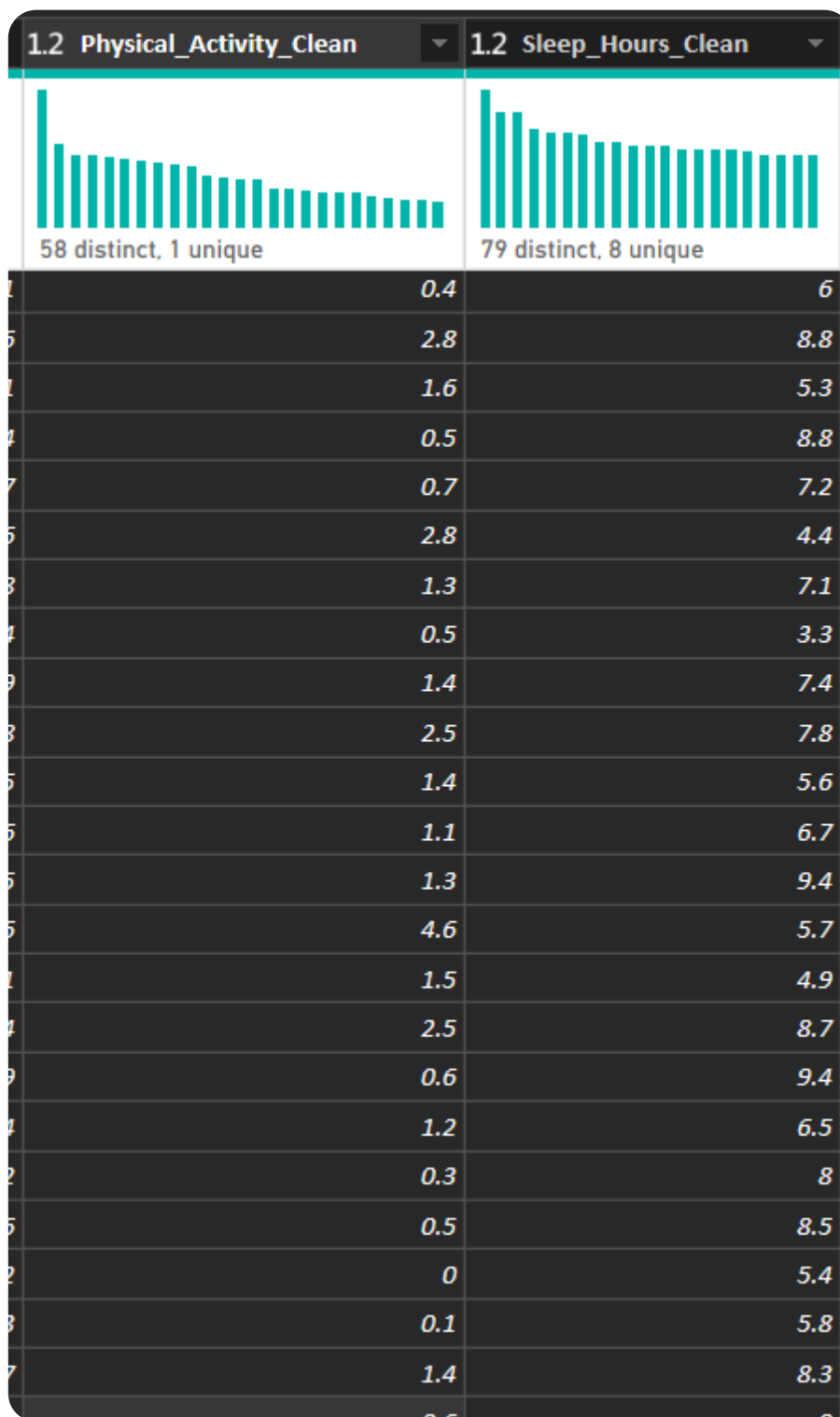
```
Columns","Sleep_Hours")),
sleepQ1 = List.Percentile(sleepList, 0.25),
sleepQ3 = List.Percentile(sleepList, 0.75),
sleepIQR = sleepQ3 - sleepQ1,
sleepLower = sleepQ1 - 1.5 * sleepIQR,
sleepUpper = sleepQ3 + 1.5 * sleepIQR,
sleepMedian = List.Median(sleepList),

// Add cleaned columns
Add_Phys_Clean = Table.AddColumn(#"Removed Columns",
"Physical_Activity_Clean",
each if [Physical_Activity_Hrs] < physLower then physMedian
else if [Physical_Activity_Hrs] > physUpper then physMedian
else [Physical_Activity_Hrs]),
Add_Sleep_Clean = Table.AddColumn(Add_Phys_Clean, "Sleep_Hours_Clean",
each if [Sleep_Hours] < sleepLower then sleepMedian
else if [Sleep_Hours] > sleepUpper then sleepMedian
else [Sleep_Hours]),
```

7. Column Management

- Removed original outlier-prone columns (Physical_Activity_Hrs, Sleep_Hours)
- Kept only the cleaned versions

Work_Stress	Self_Esteem_Score	Life_Satisfaction_Score	Loneliness_Score	Physical_Activity_Clean	Sleep_Hours_Clean
9 distinct, 0 unique	9 distinct, 0 unique	9 distinct, 0 unique	9 distinct, 0 unique	58 distinct, 1 unique	79 distinct, 8 unique
4	3	7	5	1	0.4	6
1	4	7	4	6	2.8	8.8
8	7	8	1	1	1.6	5.3
7	4	8	4	4	0.5	8.8
8	9	5	7	7	0.7	7.2
1	7	1	4	6	2.8	4.4
5	1	4	1	8	1.3	7.1
9	1	4	1	4	0.5	3.3
5	4	3	1	9	1.4	7.4
3	9	8	5	8	2.5	7.8
9	4	3	7	5	1.4	5.6
9	5	5	9	6	1.1	6.7
4	1	2	8	5	1.3	9.4
4	2	6	2	6	4.6	5.7
8	4	6	9	1	1.5	4.9
7	4	8	4	4	2.5	8.7
5	6	8	5	9	0.6	9.4
6	6	1	5	4	1.2	6.5
3	8	5	4	2	0.3	8
6	9	2	6	6	0.5	8.5
8	1	5	1	2	0	5.4
9	1	9	2	3	0.1	5.8
3	6	4	4	7	1.4	8.3
8	4	8	4	4	0.6	9

1.2 Physical_Activity_Clean ▾	1.2 Sleep_Hours_Clean ▾
58 distinct, 1 unique	79 distinct, 8 unique
0.4	6
2.8	8.8
1.6	5.3
0.5	8.8
0.7	7.2
2.8	4.4
1.3	7.1
0.5	3.3
1.4	7.4
2.5	7.8
1.4	5.6
1.1	6.7
1.3	9.4
4.6	5.7
1.5	4.9
2.5	8.7
0.6	9.4
1.2	6.5
0.3	8
0.5	8.5
0	5.4
0.1	5.8
1.4	8.3

8. Skewness Correction

- Applied log transformation to Physical_Activity_Clean using `Number.Log(value + 1)`
- Rounded the log-transformed values to 1 decimal place

```
= Table.AddColumn(#"Removed Columns1", "Custom", each Number.Log([Physical_Activity_Clean] + 1))
```

ABC 123 Custom ▼
0.336472237
1.335001067
0.955511445
0.405465108
0.530628251
1.335001067
0.832909123
0.405465108
0.875468737
1.252762968
0.875468737
0.741937345
0.832909123
1.722766598
0.916290732
1.252762968
0.470003629
0.78845736
0.262364264
0.405465108
0
0.09531018
0.875468737
0.470003629

9. Final Cleanup

- Removed original Physical_Activity_Clean
- Renamed columns for clarity:
  - "Sleep_Hours_Clean" → "Sleep_Hours"
  - "Custom" (log-transformed) → "Physical_Activity_Hours"

1.2 Sleep_Hours_Clean	▾	ABC 123 Custom	▾

↓ Sort Ascending

↓ Sort Descending

Clear Sort

Clear Filter

Remove Empty

Number Filters ▸

Search

- ☑ 0.33647223662121289
- ☑ 0.40546510810816438
- ☑ 0.47000362924573563
- ☑ 0.53062825106217038
- ☑ 0.58778666490211906
- ☑ 0.64185388617239469
- ☑ 0.69314718055994529
- ☑ 0.74193734472937733
- ☑ 0.78845736036427028
- ☑ 0.83290912293510388
- ☑ 0.87546873735389985
- ☑ 0.91629073187415511
- ☑ 0.95551144502743635
- ☑ 0.99325177301028345
- ☑ 1.0296194171811581
- ☑ 1.0647107369924282
- ☑ 1.0986122886681098
- ☑ 1.1314021114911006

⚠ List may be incomplete.   Load more

OK   Cancel

| | 1.2 Sleep_Hours_Clean | ▼ | 1.2 Custom | ▼ |

A↓ Sort Ascending

Z↓ Sort Descending

Clear Sort

▽ₓ Clear Filter

Remove Empty

Number Filters ▸

Search

☑ 0.2
☑ 0.3
☑ 0.4
☑ 0.5
☑ 0.6
☑ 0.7
☑ 0.8
☑ 0.9
☑ 1
☑ 1.1
☑ 1.2
☑ 1.3
☑ 1.4
☑ 1.5
☑ 1.6
☑ 1.7
☑ 1.8
☑ 1.9

⚠ List may be incomplete.   Load more

OK   Cancel