**Q 1: Racial disparity in ability to pay for care**

| Table of DEM\_RACE by ACC\_HCDELAY | | | |
| --- | --- | --- | --- |
| DEM\_RACE(Race/ethnicity group) | ACC\_HCDELAY(Last year ever delay in care due to cost) | | |
| Frequency Percent Row Pct Col Pct | 1:Yes | 2:No | Total |
| 1:Non-Hispanic white | 353 4.04 4.48 85.27 | 7519 86.01 95.52 90.29 | 7872 90.05 |
| 2:Non-Hispanic black | 61 0.70 7.01 14.73 | 809 9.25 92.99 9.71 | 870 9.95 |
| Total | 414 4.74 | 8328 95.26 | 8742 100.00 |

**Statistics for Table of DEM\_RACE by ACC\_HCDELAY**

| Statistic | DF | Value | Prob |
| --- | --- | --- | --- |
| Chi-Square | 1 | 11.0910 | 0.0009 |
| Likelihood Ratio Chi-Square | 1 | 9.8757 | 0.0017 |
| Continuity Adj. Chi-Square | 1 | 10.5379 | 0.0012 |
| Mantel-Haenszel Chi-Square | 1 | 11.0897 | 0.0009 |
| Phi Coefficient |  | -0.0356 |  |
| Contingency Coefficient |  | 0.0356 |  |
| Cramer's V |  | -0.0356 |  |

| Fisher's Exact Test | |
| --- | --- |
| **Cell (1,1) Frequency (F)** | 353 |
| **Left-sided Pr <= F** | 0.0010 |
| **Right-sided Pr >= F** | 0.9994 |
|  |  |
| **Table Probability (P)** | 0.0004 |
| **Two-sided Pr <= P** | 0.0018 |

**Sample Size = 8742**

From the results, we can see that a racial disparity in financial difficulties related to accessing care exists. Specifically, 4.48% of Non‐Hispanic white individuals reported delaying care due to cost, while 7.01% of Non‐Hispanic black individuals did so. Although the absolute difference between these percentages might seem modest, statistical testing confirms that it is significant. Using Fisher’s Exact Test, we obtain a two-sided p-value of 0.0018, which is well below the standard significance level of 0.05. This low p-value suggests strong evidence of a statistically significant disparity between the two groups.

**Q 2: Gender differentials in healthcare utilization**

Gender differentials in healthcare utilization:Males

| Analysis Variable : Weighted\_Visits | | |
| --- | --- | --- |
| N | Mean | Sum |
| 4520 | 3.9754425 | 17969.00 |

Gender differentials in healthcare utilization:Females

| Analysis Variable : Weighted\_Visits | | |
| --- | --- | --- |
| N | Mean | Sum |
| 5919 | 3.9216084 | 23212.00 |

From the analysis, we can see that the mean number of weighted visits is very similar for both groups—approximately 3.98 for males and 3.92 for females. This very small difference does not provide strong evidence to support the claim that one group is “lazier” or “ignorant” in seeking healthcare. We investigate this further by TTEST

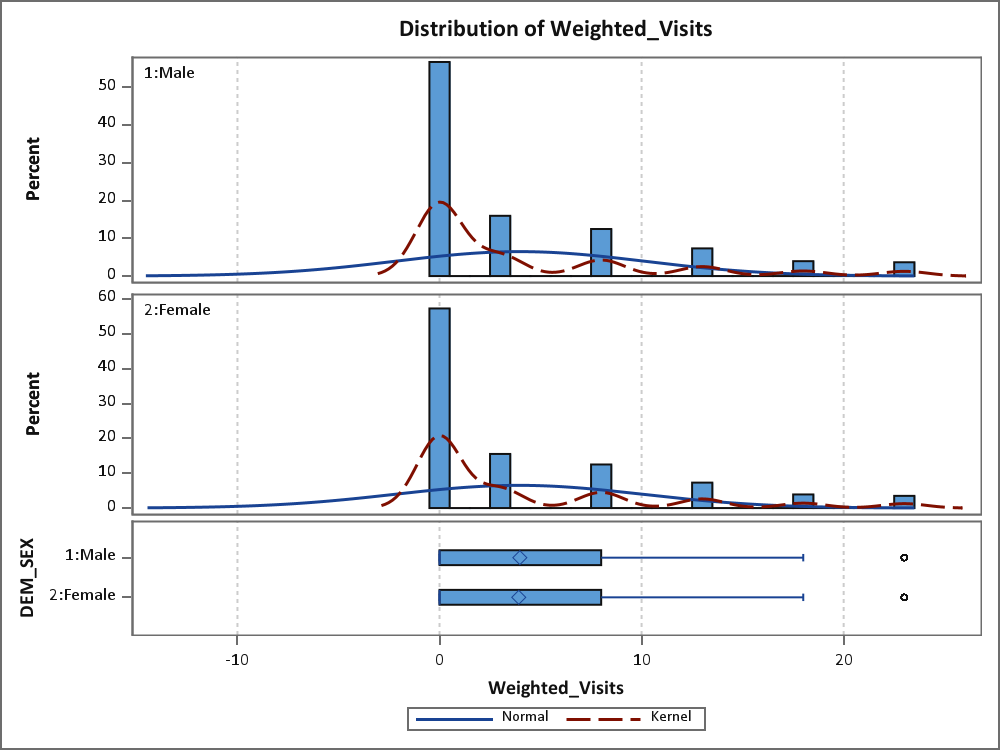
| DEM\_SEX | Method | N | Mean | Std Dev | Std Err | Minimum | Maximum |
| --- | --- | --- | --- | --- | --- | --- | --- |
| 1:Male |  | 4520 | 3.9754 | 6.1691 | 0.0918 | 0 | 23.0000 |
| 2:Female |  | 5919 | 3.9216 | 6.1242 | 0.0796 | 0 | 23.0000 |
| Diff (1-2) | Pooled |  | 0.0538 | 6.1437 | 0.1214 |  |  |
| Diff (1-2) | Satterthwaite |  | 0.0538 |  | 0.1215 |  |  |

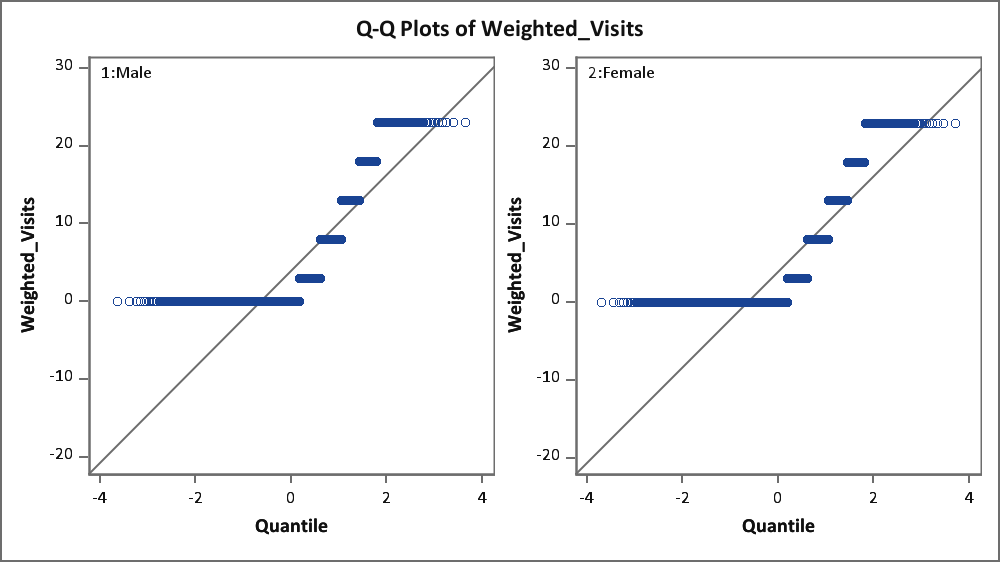
| DEM\_SEX | Method | Mean | 95% CL Mean | | Std Dev | 95% CL Std Dev | |
| --- | --- | --- | --- | --- | --- | --- | --- |
| 1:Male |  | 3.9754 | 3.7955 | 4.1553 | 6.1691 | 6.0445 | 6.2990 |
| 2:Female |  | 3.9216 | 3.7656 | 4.0777 | 6.1242 | 6.0158 | 6.2365 |
| Diff (1-2) | Pooled | 0.0538 | -0.1840 | 0.2917 | 6.1437 | 6.0614 | 6.2282 |
| Diff (1-2) | Satterthwaite | 0.0538 | -0.1843 | 0.2920 |  |  |  |

| Method | Variances | DF | t Value | Pr >  |t| |
| --- | --- | --- | --- | --- |
| Pooled | Equal | 10437 | 0.44 | 0.6573 |
| Satterthwaite | Unequal | 9689.6 | 0.44 | 0.6577 |

| Equality of Variances | | | | |
| --- | --- | --- | --- | --- |
| Method | Num  DF | Den DF | F Value | Pr > F |
| Folded F | 4519 | 5918 | 1.01 | 0.6000 |

The t-test yields a t-value of 0.44 with a p-value of about 0.657, indicating that the difference between the groups is not statistically significant. Besides, the 95% confidence intervals for both groups overlap substantially, further supporting the lack of a significant difference. An F-test for equality of variances confirmed that the variances are similar (F = 1.01, p = 0.600), which validates the use of the t-test assumptions. These findings suggest that there is no meaningful difference in the average number of weighted visits between males and females.





**Q 3: A test of relationship between education and health**

| Table of BMI\_NEW by EDU\_NEW | | | |
| --- | --- | --- | --- |
| BMI\_NEW | EDU\_NEW | | |
| Frequency Percent Row Pct Col Pct | 1 | 2 | Total |
| 1 | 3487 28.62 42.36 63.77 | 4745 38.94 57.64 70.64 | 8232 67.56 |
| 2 | 1981 16.26 50.11 36.23 | 1972 16.18 49.89 29.36 | 3953 32.44 |
| Total | 5468 44.87 | 6717 55.13 | 12185 100.00 |
| Frequency Missing = 292 | | | |

**Statistics for Table of BMI\_NEW by EDU\_NEW**

| Statistic | DF | Value | Prob |
| --- | --- | --- | --- |
| Chi-Square | 1 | 64.9217 | <.0001 |
| Likelihood Ratio Chi-Square | 1 | 64.7697 | <.0001 |
| Continuity Adj. Chi-Square | 1 | 64.6085 | <.0001 |
| Mantel-Haenszel Chi-Square | 1 | 64.9163 | <.0001 |
| Phi Coefficient |  | -0.0730 |  |
| Contingency Coefficient |  | 0.0728 |  |
| Cramer's V |  | -0.0730 |  |

| Fisher's Exact Test | |
| --- | --- |
| **Cell (1,1) Frequency (F)** | 3487 |
| **Left-sided Pr <= F** | <.0001 |
| **Right-sided Pr >= F** | 1.0000 |
|  |  |
| **Table Probability (P)** | <.0001 |
| **Two-sided Pr <= P** | <.0001 |

**Sample Size = 12185  
Frequency Missing = 292**

Less educated (EDU\_NEW = 1):

36.23% of individuals fall into the high obesity/poor health category (BMI\_NEW = 2).

Highly educated (EDU\_NEW = 2):

29.36% of individuals fall into the high obesity/poor health category.

This means that individuals with higher education are less likely to be in the high obesity/poor health category compared to those with less education. This finding is statistically significant (Fisher’s Exact Test p < .0001, along with other chi-square tests).

Regarding Folland et al.'s Theory:

Folland et al. argue that education leads to healthier lifestyles and better-informed decisions, which in turn improve health outcomes. Our results align with this theory since the highly educated group has a lower obesity rate. However, the effect size is very small (phi = –0.073), meaning that although there is a statistically significant association, education explains only a small part of the variability in obesity outcomes.

**Q 4: Are obesity and anxiety/depression related?**

| Table of BMI\_N by HLT\_N | | | |
| --- | --- | --- | --- |
| BMI\_N | HLT\_N | | |
| Frequency Percent Row Pct Col Pct | 1 | 2 | Total |
| 1 | 1890 15.49 22.95 57.92 | 6347 52.01 77.05 70.99 | 8237 67.49 |
| 2 | 1373 11.25 34.61 42.08 | 2594 21.26 65.39 29.01 | 3967 32.51 |
| Total | 3263 26.74 | 8941 73.26 | 12204 100.00 |
| Frequency Missing = 273 | | | |

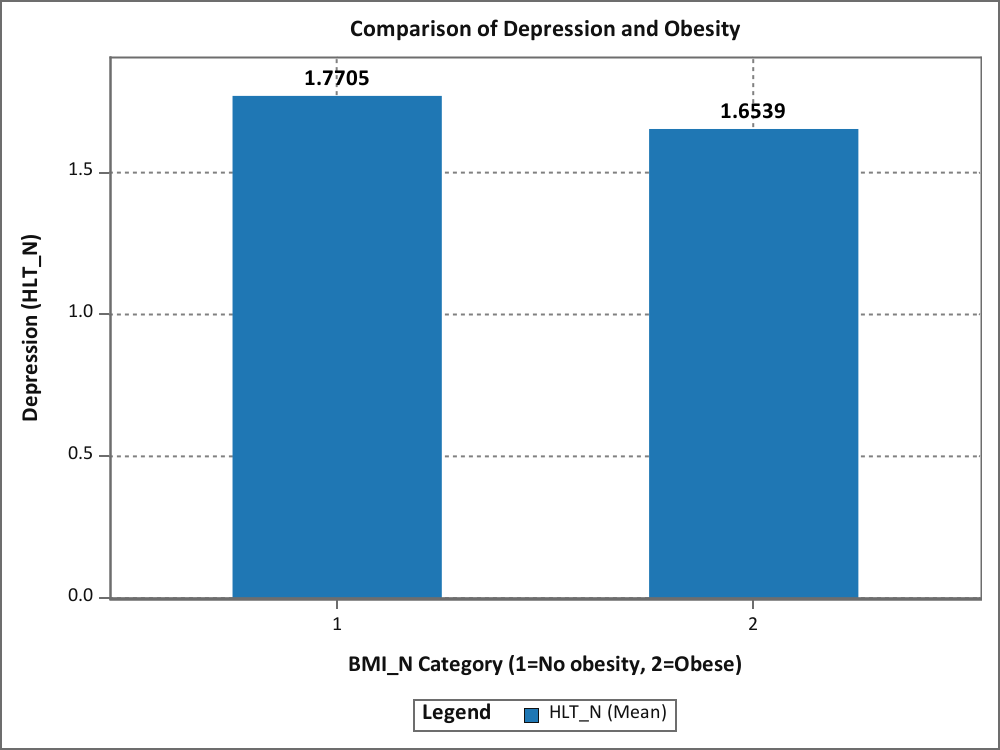
**Statistics for Table of BMI\_N by HLT\_N**

| Statistic | DF | Value | Prob |
| --- | --- | --- | --- |
| Chi-Square | 1 | 186.0036 | <.0001 |
| Likelihood Ratio Chi-Square | 1 | 181.3642 | <.0001 |
| Continuity Adj. Chi-Square | 1 | 185.4086 | <.0001 |
| Mantel-Haenszel Chi-Square | 1 | 185.9884 | <.0001 |
| Phi Coefficient |  | -0.1235 |  |
| Contingency Coefficient |  | 0.1225 |  |
| Cramer's V |  | -0.1235 |  |

| Fisher's Exact Test | |
| --- | --- |
| **Cell (1,1) Frequency (F)** | 1890 |
| **Left-sided Pr <= F** | <.0001 |
| **Right-sided Pr >= F** | 1.0000 |
|  |  |
| **Table Probability (P)** | <.0001 |
| **Two-sided Pr <= P** | <.0001 |

**Sample Size = 12204  
Frequency Missing = 273**

Among individuals classified as healthy (BMI\_N = 1) and obese (BMI\_N = 2), the depression rates are 22.95% and 34.61%, respectively. A Fisher’s Exact Test for this 2×2 table yields a two-sided p-value of < .0001, indicating that the difference in depression rates between the two groups is statistically significant and suggesting that obese individuals are more likely to experience depression compared to their healthy counterparts.



**Q 5: Time to make an educated guess and test it!**

A screenshot of a computer

AI-generated content may be incorrect.A screenshot of a table

AI-generated content may be incorrect.A screenshot of a data

AI-generated content may be incorrect.A screenshot of a table

AI-generated content may be incorrect.

The above analysis shows the relationship between depression(HLT\_OCDEPRSS) and obesity(HLT\_BMI\_CAT) categorized by Gender(DEM\_SEX=2).

Males, sample size is 5,519 and 60 missing observations, depression increases with BMI: category 3 is obese male and they account for 29.32 percent, while overweight(2) males account for 17.68, and normal ones account for 19.73 . A Chi-Square test (p < 0.0001) shows a significant relationship, with a weak association (Phi coefficient = 0.1123).

Females, 6,685 samples and 178 missing observations, the depression rate is higher for them: normal person account for 23.44 %, overweight people are 30.72 %, and 39 percent of them are obese. Chi-Square results (p < 0.0001) indicate a highly significant relationship, with a stronger association (Phi coefficient = 0.1447).

In summary, obesity is linked to higher depression rates in both genders, with a stronger association in females. However, other factors beyond BMI may also contribute, and the higher missing data in females could suggest reporting biases.

**Q 6a: Loneliness and health**

| Table of HELTH\_NEW by MARSTA\_NEW | | | |
| --- | --- | --- | --- |
| HELTH\_NEW | MARSTA\_NEW | | |
| Frequency Percent Row Pct Col Pct | 1 | 2 | Total |
| 1 | 2626 39.00 51.93 77.67 | 2431 36.11 48.07 72.52 | 5057 75.11 |
| 2 | 755 11.21 45.05 22.33 | 921 13.68 54.95 27.48 | 1676 24.89 |
| Total | 3381 50.22 | 3352 49.78 | 6733 100.00 |
| Frequency Missing = 3706 | | | |

**Statistics for Table of HELTH\_NEW by MARSTA\_NEW**

| Statistic | DF | Value | Prob |
| --- | --- | --- | --- |
| Chi-Square | 1 | 23.8363 | <.0001 |
| Likelihood Ratio Chi-Square | 1 | 23.8648 | <.0001 |
| Continuity Adj. Chi-Square | 1 | 23.5619 | <.0001 |
| Mantel-Haenszel Chi-Square | 1 | 23.8328 | <.0001 |
| Phi Coefficient |  | 0.0595 |  |
| Contingency Coefficient |  | 0.0594 |  |
| Cramer's V |  | 0.0595 |  |

| Fisher's Exact Test | |
| --- | --- |
| **Cell (1,1) Frequency (F)** | 2626 |
| **Left-sided Pr <= F** | 1.0000 |
| **Right-sided Pr >= F** | <.0001 |
|  |  |
| **Table Probability (P)** | <.0001 |
| **Two-sided Pr <= P** | <.0001 |

**Sample Size = 6733  
Frequency Missing = 3706**

There is a statistically significant association between living arrangements and health status, with those living with family/partners more likely to report good health (77.67%) compared to those living alone (72.52%). However, despite the significance (p < .0001), the strength of this correlation is very weak, as indicated by a Cramer’s V of approximately 0.06.

**Q 6b: Loneliness and risk of depression**

| Table of HLT\_NEW by MARSTA\_NEW | | | |
| --- | --- | --- | --- |
| HLT\_NEW | MARSTA\_NEW | | |
| Frequency Percent Row Pct Col Pct | 1 | 2 | Total |
| 1 | 949 9.12 42.92 17.76 | 1262 12.13 57.08 24.93 | 2211 21.25 |
| 2 | 4393 42.22 53.62 82.24 | 3800 36.52 46.38 75.07 | 8193 78.75 |
| Total | 5342 51.35 | 5062 48.65 | 10404 100.00 |
| Frequency Missing = 35 | | | |

**Statistics for Table of HLT\_NEW by MARSTA\_NEW**

| Statistic | DF | Value | Prob |
| --- | --- | --- | --- |
| Chi-Square | 1 | 79.7527 | <.0001 |
| Likelihood Ratio Chi-Square | 1 | 79.8808 | <.0001 |
| Continuity Adj. Chi-Square | 1 | 79.3251 | <.0001 |
| Mantel-Haenszel Chi-Square | 1 | 79.7450 | <.0001 |
| Phi Coefficient |  | -0.0876 |  |
| Contingency Coefficient |  | 0.0872 |  |
| Cramer's V |  | -0.0876 |  |

| Fisher's Exact Test | |
| --- | --- |
| **Cell (1,1) Frequency (F)** | 949 |
| **Left-sided Pr <= F** | <.0001 |
| **Right-sided Pr >= F** | 1.0000 |
|  |  |
| **Table Probability (P)** | <.0001 |
| **Two-sided Pr <= P** | <.0001 |

**Sample Size = 10404  
Frequency Missing = 35**

The analysis indicates that among individuals with depression (HLT\_NEW = 1), 57.08% are living alone, whereas among those without depression (HLT\_NEW = 2), only 46.38% live alone, suggesting a slightly higher risk of depression for those living alone. This shows a statistically significant association (p < .0001 across all Chi-Square tests), however the effect size is very small (phi ≈ -0.0876), indicating that the practical impact of living arrangement on depression risk is minimal.