**The FREQ Procedure**

| Diabetes_01 | Frequency | Percent | Cumulative Frequency | Cumulative Percent |
|---|---|---|---|---|
| 0 | 213703 | 84.24 | 213703 | 84.24 |
| 1 | 39977 | 15.76 | 253680 | 100.00 |
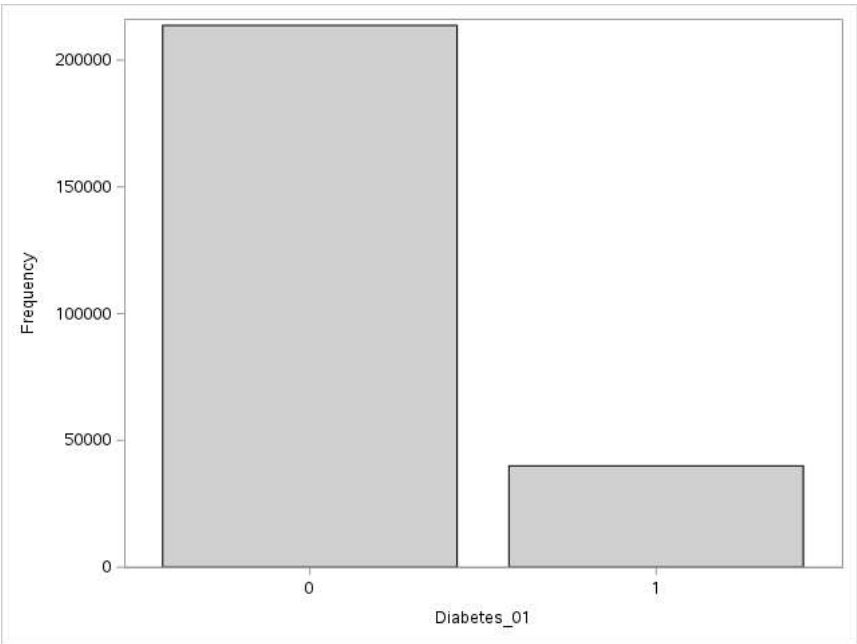


**The FREQ Procedure**

| Age | Frequency | Percent | Cumulative Frequency | Cumulative Percent |
|---|---|---|---|---|
| 1 | 5700 | 2.25 | 5700 | 2.25 |
| 2 | 7598 | 3.00 | 13298 | 5.24 |
| 3 | 11123 | 4.38 | 24421 | 9.63 |
| 4 | 13823 | 5.45 | 38244 | 15.08 |
| 5 | 16157 | 6.37 | 54401 | 21.44 |
| 6 | 19819 | 7.81 | 74220 | 29.26 |
| 7 | 26314 | 10.37 | 100534 | 39.63 |
| 8 | 30832 | 12.15 | 131366 | 51.78 |
| 9 | 33244 | 13.10 | 164610 | 64.89 |
| 10 | 32194 | 12.69 | 196804 | 77.58 |
| 11 | 23533 | 9.28 | 220337 | 86.86 |
| 12 | 15980 | 6.30 | 236317 | 93.16 |
| 13 | 17363 | 6.84 | 253680 | 100.00 |
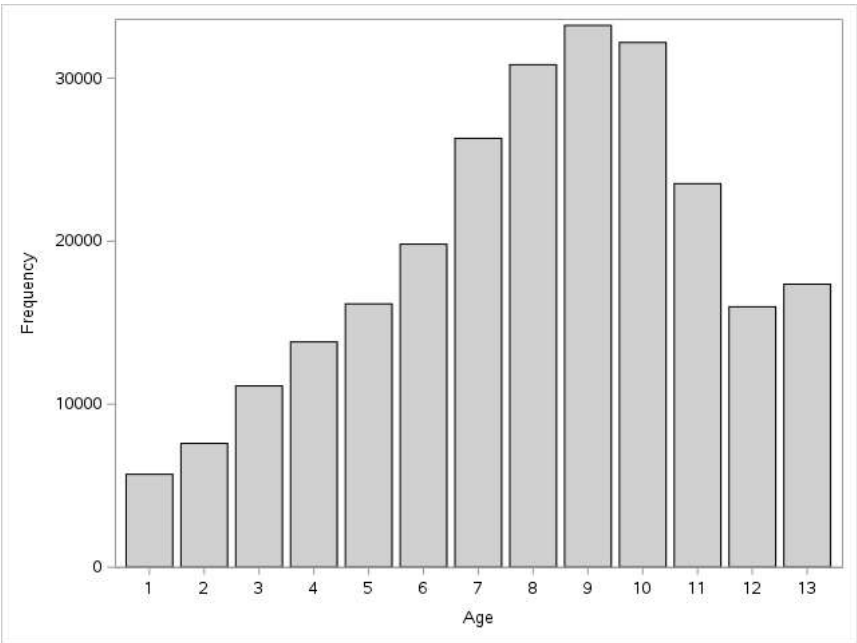


**The FREQ Procedure**

| Age | AgeGroup | Frequency | Percent | Cumulative Frequency | Cumulative Percent |
|---|---|---|---|---|---|
| 1 | Under 40 | 5700 | 2.25 | 5700 | 2.25 |
| 2 | Under 40 | 7598 | 3.00 | 13298 | 5.24 |
| 3 | Under 40 | 11123 | 4.38 | 24421 | 9.63 |
| 4 | Under 40 | 13823 | 5.45 | 38244 | 15.08 |
| 5 | Under 40 | 16157 | 6.37 | 54401 | 21.44 |
| 6 | 40-59 | 19819 | 7.81 | 74220 | 29.26 |
| 7 | 40-59 | 26314 | 10.37 | 100534 | 39.63 |
| 8 | 40-59 | 30832 | 12.15 | 131366 | 51.78 |
| 9 | 60+ | 33244 | 13.10 | 164610 | 64.89 |
| 10 | 60+ | 32194 | 12.69 | 196804 | 77.58 |
| 11 | 60+ | 23533 | 9.28 | 220337 | 86.86 |
| 12 | 60+ | 15980 | 6.30 | 236317 | 93.16 |
| 13 | 60+ | 17363 | 6.84 | 253680 | 100.00 |

| AgeGroup | Frequency | Percent | Cumulative Frequency | Cumulative Percent |
|---|---|---|---|---|
| 40-59 | 76965 | 30.34 | 76965 | 30.34 |
| 60+ | 122314 | 48.22 | 199279 | 78.56 |
| Under 40 | 54401 | 21.44 | 253680 | 100.00 |
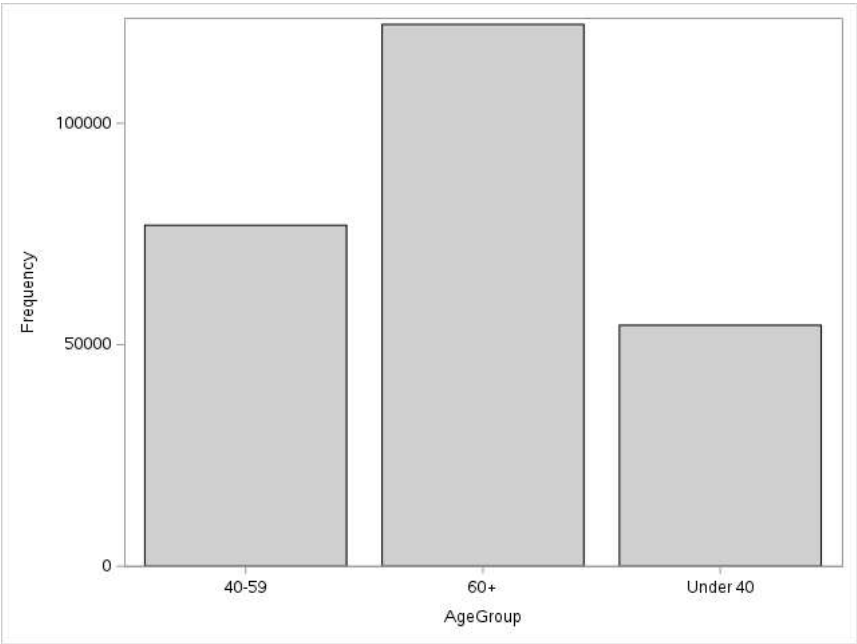
### The UNIVARIATE Procedure
### Variable: BMI

| Moments | | | |
|---|---|---|---|
| N | 253680 | Sum Weights | 253680 |
| Mean | 28.3823636 | Sum Observations | 7200038 |
| Std Deviation | 6.6086942 | Variance | 43.674839 |
| Skewness | 2.12200376 | Kurtosis | 10.9974733 |
| Uncorrected SS | 215433486 | Corrected SS | 11079389.5 |
| Coeff Variation | 23.2845097 | Std Error Mean | 0.01312117 |

| Basic Statistical Measures | | | |
|---|---|---|---|
| Location | | Variability | |
| Mean | 28.38236 | Std Deviation | 6.60869 |
| Median | 27.00000 | Variance | 43.67484 |
| Mode | 27.00000 | Range | 86.00000 |
| | | Interquartile Range | 7.00000 |

| Tests for Location: Mu0=0 | | | | |
|---|---|---|---|---|
| Test | Statistic | | p Value | |
| Student's t | t | 2163.097 | Pr > \|t\| | <.0001 |
| Sign | M | 126840 | Pr >= \|M\| | <.0001 |
| Signed Rank | S | 1.609E10 | Pr >= \|S\| | <.0001 |

| Quantiles (Definition 5) | |
|---|---|
| Level | Quantile |
| 100% Max | 98 |
| 99% | 50 |
| 95% | 40 |
| 90% | 36 |
| 75% Q3 | 31 |
| 50% Median | 27 |
| 25% Q1 | 24 |
| 10% | 22 |
| 5% | 20 |
| 1% | 18 |
| 0% Min | 12 |

| Extreme Observations | | | |
|---|---|---|---|
| Lowest | | Highest | |
| Value | Obs | Value | Obs |
| 12 | 239608 | 98 | 76397 |
| 12 | 210864 | 98 | 76533 |
| 12 | 177463 | 98 | 79479 |
| 12 | 136037 | 98 | 79508 |
| 12 | 98003 | 98 | 79555 |

### The UNIVARIATE Procedure
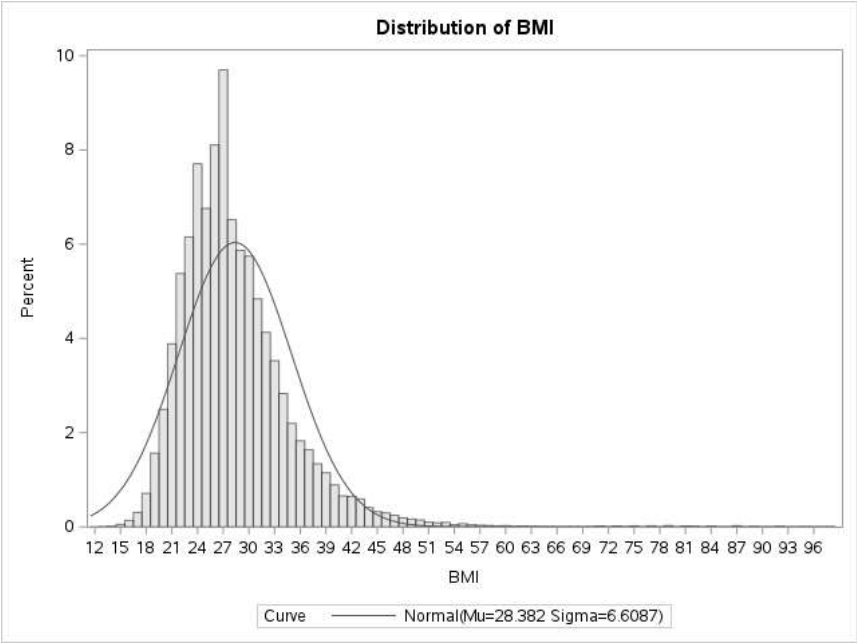


### The UNIVARIATE Procedure
### Fitted Normal Distribution for BMI

| Parameters for Normal Distribution | | |
|---|---|---|
| Parameter | Symbol | Estimate |
| Mean | Mu | 28.38236 |
| Std Dev | Sigma | 6.608694 |

| Goodness-of-Fit Tests for Normal Distribution | | | | |
|---|---|---|---|---|
| Test | Statistic | | p Value | |
| Kolmogorov-Smirnov | D | 0.11807 | Pr > D | <0.010 |
| Cramer-von Mises | W-Sq | 867.60802 | Pr > W-Sq | <0.005 |
| Anderson-Darling | A-Sq | 5128.22547 | Pr > A-Sq | <0.005 |

| Quantiles for Normal Distribution | | |
|---|---|---|
| | Quantile | |
| Percent | Observed | Estimated |
| 1.0 | 18.0000 | 13.0082 |
| 5.0 | 20.0000 | 17.5120 |
| 10.0 | 22.0000 | 19.9130 |

| Quantiles for Normal Distribution | | |
|---|---|---|
| | Quantile | |
| Percent | Observed | Estimated |
| 25.0 | 24.0000 | 23.9249 |
| 50.0 | 27.0000 | 28.3824 |
| 75.0 | 31.0000 | 32.8399 |
| 90.0 | 36.0000 | 36.8517 |
| 95.0 | 40.0000 | 39.2527 |
| 99.0 | 50.0000 | 43.7565 |

The UNIVARIATE Procedure



Probability Plot for BMI

Normal Line —— Mu=28.382, Sigma=6.6087
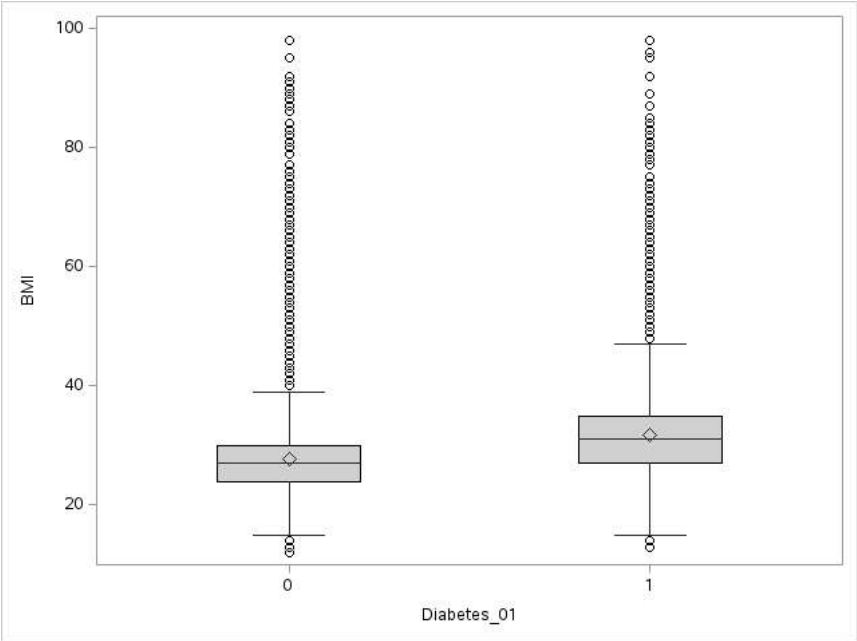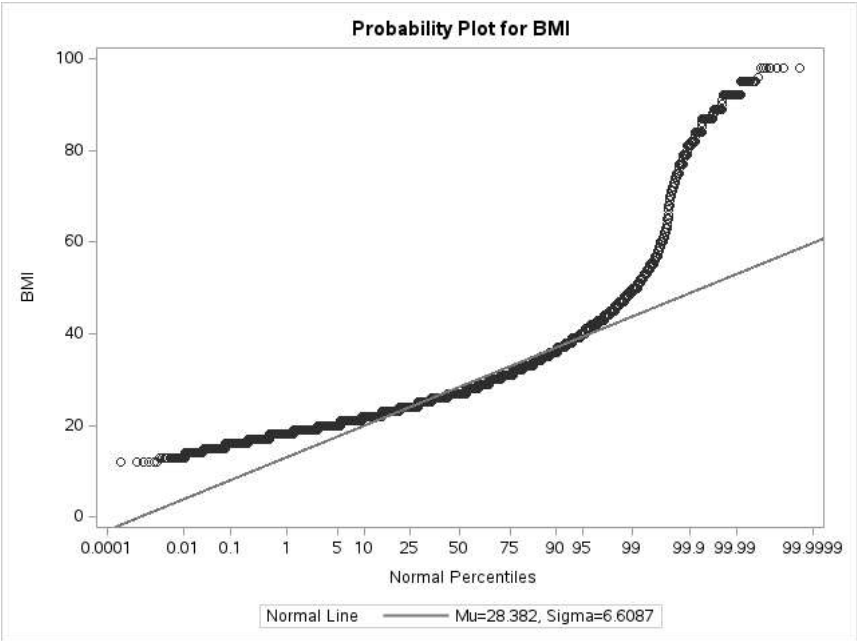


The ANOVA Procedure

| Class Level Information | | |
|---|---|---|
| Class | Levels | Values |
| AgeGroup | 3 | 40-59 60+ Under 40 |

| Number of Observations Read | 253680 |
|---|---|
| Number of Observations Used | 253680 |

The ANOVA Procedure

Dependent Variable: BMI

| Source | DF | Sum of Squares | Mean Square | F Value | Pr > F |
|---|---|---|---|---|---|
| Model | 2 | 36663.80 | 18331.90 | 421.13 | <.0001 |
| Error | 253677 | 11042725.70 | 43.53 | | |
| Corrected Total | 253679 | 11079389.49 | | | |

| R-Square | Coeff Var | Root MSE | BMI Mean |
|---|---|---|---|
| 0.003309 | 23.24604 | 6.597776 | 28.38236 |

| Source | DF | Anova SS | Mean Square | F Value | Pr > F |
|---|---|---|---|---|---|
| AgeGroup | 2 | 36663.79882 | 18331.89941 | 421.13 | <.0001 |

The ANOVA Procedure



Distribution of BMI
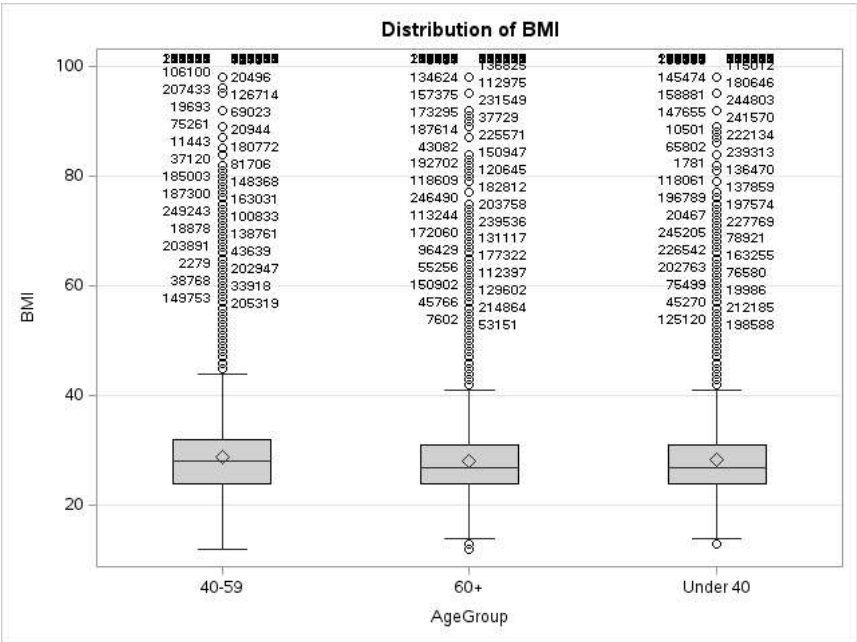
### The ANOVA Procedure

#### Tukey's Studentized Range (HSD) Test for BMI

**Note:** This test controls the Type I experimentwise error rate.

| | |
|---|---|
| Alpha | 0.05 |
| Error Degrees of Freedom | 253677 |
| Error Mean Square | 43.53065 |
| Critical Value of Studentized Range | 3.31451 |

**Comparisons significant at the 0.05 level are indicated by \*\*\*.**

| AgeGroup Comparison | Difference Between Means | Simultaneous 95% Confidence Limits | | |
|---|---|---|---|---|
| 40-59 - Under 40 | 0.69714 | 0.61052 | 0.78375 | \*\*\* |
| 40-59 - 60+ | 0.86688 | 0.79573 | 0.93802 | \*\*\* |
| Under 40 - 40-59 | -0.69714 | -0.78375 | -0.61052 | \*\*\* |
| Under 40 - 60+ | 0.16974 | 0.09005 | 0.24943 | \*\*\* |
| 60+ - 40-59 | -0.86688 | -0.93802 | -0.79573 | \*\*\* |
| 60+ - Under 40 | -0.16974 | -0.24943 | -0.09005 | \*\*\* |

---

### The FASTCLUS Procedure
#### Replace=FULL Radius=0 Maxclusters=3 Maxiter=1

**Initial Seeds**

| Cluster | HighBP | HighChol | CholCheck | Smoker | Stroke | HeartDiseaseorAttack | PhysActivity | Fruits | Veggies | HvyAlcoholConsump | AnyHealthcare | NoDocbcCost | GenHlth | MentHlth | BMI | PhysHlth | DiffWalk | Age | Education |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 1.00000000 | 1.00000000 | 1.00000000 | 0.00000000 | 1.00000000 | 0.00000000 | 0.00000000 | 0.00000000 | 1.00000000 | 0.00000000 | 1.00000000 | 0.00000000 | 2.00000000 | 0.00000000 | 98.00000000 | 0.00000000 | 0.00000000 | 13.00000000 | 5.00000000 |
| 2 | 0.00000000 | 0.00000000 | 1.00000000 | 0.00000000 | 0.00000000 | 0.00000000 | 1.00000000 | 1.00000000 | 1.00000000 | 0.00000000 | 0.00000000 | 0.00000000 | 3.00000000 | 12.00000000 | 0.00000000 | 0.00000000 | 1.00000000 | 13.00000000 | 5.00000000 |
| 3 | 0.00000000 | 1.00000000 | 1.00000000 | 1.00000000 | 0.00000000 | 1.00000000 | 1.00000000 | 0.00000000 | 1.00000000 | 0.00000000 | 1.00000000 | 0.00000000 | 5.00000000 | 30.00000000 | 56.00000000 | 30.00000000 | 1.00000000 | 1.00000000 | 4.00000000 |

| | |
|---|---|
| Criterion Based on Final Seeds = | 2.4594 |

**Cluster Summary**

| Cluster | Frequency | RMS Std Deviation | Maximum Distance from Seed to Observation | Radius Exceeded | Nearest Cluster | Distance Between Cluster Centroids |
|---|---|---|---|---|---|---|
| 1 | 1413 | 3.8142 | 45.5717 | | 2 | 38.9955 |
| 2 | 220065 | 1.9327 | 31.4190 | | 3 | 25.5031 |
| 3 | 32202 | 4.2329 | 39.1357 | | 2 | 25.5031 |

**Statistics for Variables**

| Variable | Total STD | Within STD | R-Square | RSQ/(1-RSQ) |
|---|---|---|---|---|
| HighBP | 0.49493 | 0.49035 | 0.018460 | 0.018807 |
| HighChol | 0.49421 | 0.49149 | 0.010972 | 0.011094 |
| CholCheck | 0.18957 | 0.18954 | 0.000350 | 0.000350 |
| Smoker | 0.49676 | 0.49351 | 0.013045 | 0.013217 |
| Stroke | 0.19729 | 0.19572 | 0.015924 | 0.016182 |
| HeartDiseaseorAttack | 0.29209 | 0.28881 | 0.022323 | 0.022832 |
| PhysActivity | 0.42917 | 0.42023 | 0.041213 | 0.042984 |
| Fruits | 0.48164 | 0.48082 | 0.003425 | 0.003436 |
| Veggies | 0.39118 | 0.39029 | 0.004525 | 0.004546 |
| HvyAlcoholConsump | 0.23030 | 0.23027 | 0.000267 | 0.000267 |
| AnyHealthcare | 0.21576 | 0.21568 | 0.000772 | 0.000772 |
| NoDocbcCost | 0.27765 | 0.27398 | 0.026285 | 0.026995 |
| GenHlth | 1.06848 | 0.94932 | 0.210604 | 0.266791 |
| MentHlth | 7.41285 | 5.88930 | 0.368820 | 0.584333 |
| BMI | 6.60869 | 5.88136 | 0.208008 | 0.262640 |
| PhysHlth | 8.71795 | 5.07567 | 0.661036 | 1.950163 |
| DiffWalk | 0.37407 | 0.34063 | 0.170806 | 0.205990 |
| Age | 3.05422 | 3.04941 | 0.003157 | 0.003167 |
| Education | 0.98577 | 0.97457 | 0.022616 | 0.023140 |
| Income | 2.07115 | 1.99960 | 0.067907 | 0.072854 |
| OVER-ALL | 3.09952 | 2.36557 | 0.417522 | 0.716803 |

| | |
|---|---|
| Pseudo F Statistic = | 90918.24 |

| | |
|---|---|
| Approximate Expected Over-All R-Squared = | 0.45733 |

| | |
|---|---|
| Cubic Clustering Criterion = | -90.921 |

**WARNING: The two values above are invalid for correlated variables.**

**Cluster Means**

| Cluster | HighBP | HighChol | CholCheck | Smoker | Stroke | HeartDiseaseorAttack | PhysActivity | Fruits | Veggies | HvyAlcoholConsump | AnyHealthcare | NoDocbcCost | GenHlth | MentHlth | BMI | PhysHlth | DiffWalk | Age | Education |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 0.54493984 | 0.39561217 | 0.96319887 | 0.37154989 | 0.03255485 | 0.08634112 | 0.59377212 | 0.53290870 | 0.76008493 | 0.03113942 | 0.91861288 | 0.10898797 | 2.96461430 | 2.95753715 | 66.60934183 | 4.65534324 | 0.36730361 | 6.74734607 | 4.90728946 |
| 2 | 0.40277191 | 0.40443960 | 0.96130689 | 0.42190262 | 0.03106809 | 0.07749074 | 0.79054825 | 0.64523663 | 0.82169359 | 0.05758753 | 0.95333651 | 0.06678481 | 2.32120964 | 1.45914616 | 27.80440324 | 1.52081885 | 0.10808170 | 7.98479540 | 5.10795447 |
| 3 | 0.60316129 | 0.55987206 | 0.97195826 | 0.59164027 | 0.10586299 | 0.20862058 | 0.53130240 | 0.56366064 | 0.74346314 | 0.04779206 | 0.93686728 | 0.20194398 | 3.79119309 | 14.98748525 | 30.65471089 | 22.82075647 | 0.57049252 | 8.41189988 | 4.66362338 |

**Cluster Standard Deviations**

| Cluster | HighBP | HighChol | CholCheck | Smoker | Stroke | HeartDiseaseorAttack | PhysActivity | Fruits | Veggies | HvyAlcoholConsump | AnyHealthcare | NoDocbcCost | GenHlth | MentHlth | BMI | PhysHlth | DiffWalk | Age | Education |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 0.49815262 | 0.48915490 | 0.18833989 | 0.48339005 | 0.17753122 | 0.28096654 | 0.49130196 | 0.49909248 | 0.42718261 | 0.17375593 | 0.27352551 | 0.31173445 | 1.06323798 | 6.32508452 | 12.85312422 | 8.30027612 | 0.48224086 | 2.78105725 | 1.01645628 |
| 2 | 0.49045672 | 0.49078438 | 0.19286297 | 0.49386426 | 0.17350217 | 0.26736913 | 0.40691826 | 0.47844264 | 0.38277134 | 0.23296234 | 0.21091754 | 0.24964953 | 0.93176297 | 3.86049404 | 5.51237946 | 3.55212187 | 0.31048428 | 3.08997209 | 0.96283434 |
| 3 | 0.48924961 | 0.49641010 | 0.16509465 | 0.49153796 | 0.30766696 | 0.40632888 | 0.49902695 | 0.49593846 | 0.43672832 | 0.21332931 | 0.24320531 | 0.40145686 | 1.05684497 | 13.02425392 | 7.58906155 | 10.66319744 | 0.49501355 | 2.76867400 | 1.04950842 |

---

### The FREQ Procedure

| Diabetes_012 | CLUSTER | Frequency | Percent | Cumulative Frequency | Cumulative Percent |
|---|---|---|---|---|---|
| 0 | 1 | 1027 | 0.40 | 1027 | 0.40 |
| 0 | 2 | 190260 | 75.00 | 191287 | 75.40 |
| 0 | 3 | 22416 | 8.84 | 213703 | 84.24 |
| 1 | 1 | 39 | 0.02 | 213742 | 84.26 |
| 1 | 2 | 3645 | 1.44 | 217387 | 85.69 |
| 1 | 3 | 947 | 0.37 | 218334 | 86.07 |
| 2 | 1 | 347 | 0.14 | 218681 | 86.20 |
| 2 | 2 | 26160 | 10.31 | 244841 | 96.52 |
| 2 | 3 | 8839 | 3.48 | 253680 | 100.00 |

---

### The FASTCLUS Procedure
#### Replace=FULL Radius=0 Maxclusters=2 Maxiter=1

**Initial Seeds**

| Cluster | HighBP | HighChol | CholCheck | Smoker | Stroke | HeartDiseaseorAttack | PhysActivity | Fruits | Veggies | HvyAlcoholConsump | AnyHealthcare | NoDocbcCost | GenHlth | MentHlth | BMI | PhysHlth | DiffWalk | Age | Education |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 0.00000000 | 0.00000000 | 1.00000000 | 0.00000000 | 0.00000000 | 0.00000000 | 1.00000000 | 1.00000000 | 1.00000000 | 0.00000000 | 0.00000000 | 0.00000000 | 3.00000000 | 0.00000000 | 12.00000000 | 0.00000000 | 1.00000000 | 13.00000000 | 5.00000000 |
| 2 | 1.00000000 | 1.00000000 | 1.00000000 | 1.00000000 | 0.00000000 | 1.00000000 | 0.00000000 | 1.00000000 | 1.00000000 | 0.00000000 | 1.00000000 | 0.00000000 | 5.00000000 | 30.00000000 | 98.00000000 | 30.00000000 | 1.00000000 | 8.00000000 | 6.00000000 |

| | |
|---|---|
| Criterion Based on Final Seeds = | 2.9920 |

**Cluster Summary**

| Cluster | Frequency | RMS Std Deviation | Maximum Distance from Seed to Observation | Radius Exceeded | Nearest Cluster | Distance Between Cluster Centroids |
|---|---|---|---|---|---|---|
| 1 | 248676 | 2.8964 | 40.9859 | | 2 | 30.2350 |
| 2 | 5004 | 5.0327 | 39.0160 | | 1 | 30.2350 |

**Statistics for Variables**

| Variable | Total STD | Within STD | R-Square | RSQ/(1-RSQ) |
|---|---|---|---|---|
| HighBP | 0.49493 | 0.49361 | 0.005335 | 0.005364 |
| HighChol | 0.49421 | 0.49390 | 0.001264 | 0.001265 |
| CholCheck | 0.18957 | 0.18956 | 0.000106 | 0.000106 |

Results: diabetes_analysis1.sas

| Statistics for Variables | | | | |
|---|---|---|---|---|
| Variable | Total STD | Within STD | R-Square | RSQ/(1-RSQ) |
| Smoker | 0.49676 | 0.49668 | 0.000338 | 0.000339 |
| Stroke | 0.19729 | 0.19718 | 0.001199 | 0.001200 |
| HeartDiseaseorAttack | 0.29209 | 0.29180 | 0.001939 | 0.001943 |
| PhysActivity | 0.42917 | 0.42704 | 0.009892 | 0.009991 |
| Fruits | 0.48164 | 0.48140 | 0.000998 | 0.000999 |
| Veggies | 0.39118 | 0.39101 | 0.000838 | 0.000838 |
| HvyAlcoholConsump | 0.23030 | 0.23026 | 0.000356 | 0.000356 |
| AnyHealthcare | 0.21576 | 0.21572 | 0.000324 | 0.000324 |
| NoDocbcCost | 0.27765 | 0.27706 | 0.004294 | 0.004312 |
| GenHlth | 1.06848 | 1.05242 | 0.029841 | 0.030759 |
| MentHlth | 7.41285 | 7.23688 | 0.046917 | 0.049227 |
| BMI | 6.60869 | 5.75719 | 0.241093 | 0.317685 |
| PhysHlth | 8.71795 | 8.45858 | 0.058622 | 0.062273 |
| DiffWalk | 0.37407 | 0.36738 | 0.035415 | 0.036715 |
| Age | 3.05422 | 3.05255 | 0.001096 | 0.001097 |
| Education | 0.98577 | 0.98424 | 0.003114 | 0.003123 |
| Income | 2.07115 | 2.05742 | 0.013217 | 0.013394 |
| OVER-ALL | 3.09952 | 2.95352 | 0.091999 | 0.101320 |

| Pseudo F Statistic = | 25702.63 |
|---|---|

| Approximate Expected Over-All R-Squared = | 0.29667 |
|---|---|

| Cubic Clustering Criterion = | -389.405 |
|---|---|

WARNING: The two values above are invalid for correlated variables.

| Cluster Means | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Cluster | HighBP | HighChol | CholCheck | Smoker | Stroke | HeartDiseaseorAttack | PhysActivity | Fruits | Veggies | HvyAlcoholConsump | AnyHealthcare | NoDocbcCost | GenHlth | MentHlth | BMI | PhysHlth | DiffWalk | Age | Education |
| 1 | 0.42387283 | 0.42162895 | 0.96239283 | 0.44187215 | 0.03960173 | 0.09236114 | 0.76259872 | 0.63641445 | 0.81302578 | 0.05681288 | 0.95160369 | 0.08159613 | 2.48520967 | 2.95700429 | 27.92205520 | 3.94265631 | 0.15823803 | 8.04646206 | 5.05823642 |
| 2 | 0.68385292 | 0.54796163 | 0.97641886 | 0.50759392 | 0.08872902 | 0.18485212 | 0.45563549 | 0.52697842 | 0.73161471 | 0.02557954 | 0.92366107 | 0.21243006 | 3.81254996 | 14.50379696 | 51.25759392 | 19.12210232 | 0.66446843 | 7.31934452 | 4.66266986 |

| Cluster Standard Deviations | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Cluster | HighBP | HighChol | CholCheck | Smoker | Stroke | HeartDiseaseorAttack | PhysActivity | Fruits | Veggies | HvyAlcoholConsump | AnyHealthcare | NoDocbcCost | GenHlth | MentHlth | BMI | PhysHlth | DiffWalk | Age | Education |
| 1 | 0.49417167 | 0.49382078 | 0.19024462 | 0.49661066 | 0.19502202 | 0.28953566 | 0.42549106 | 0.48103225 | 0.38989161 | 0.23148519 | 0.21460263 | 0.27374898 | 1.05217770 | 7.07193966 | 5.52108080 | 8.35855490 | 0.36496478 | 3.06186156 | 0.98322232 |
| 2 | 0.46501755 | 0.49774410 | 0.15175529 | 0.49999229 | 0.28438063 | 0.38821635 | 0.49807767 | 0.49932153 | 0.44316348 | 0.15789302 | 0.26556617 | 0.40906841 | 1.06421560 | 13.02609604 | 12.86494804 | 12.45610639 | 0.47222315 | 2.54732098 | 1.03357456 |

The FREQ Procedure

| Diabetes_01 | CLUSTER | Frequency | Percent | Cumulative Frequency | Cumulative Percent |
|---|---|---|---|---|---|
| 0 | 1 | 210877 | 83.13 | 210877 | 83.13 |
| 0 | 2 | 2826 | 1.11 | 213703 | 84.24 |
| 1 | 1 | 37799 | 14.90 | 251502 | 99.14 |
| 1 | 2 | 2178 | 0.86 | 253680 | 100.00 |