**Table 1: Sample characteristics**

|  |  |  |  |
| --- | --- | --- | --- |
|  | **Category** | **Frequency** | **Percent** |
| **Age** | <25 | 684 | 9.52 |
|  | 26-35 | 3018 | 42.02 |
|  | 36-45 | 1919 | 26.72 |
|  | 46-55 | 1099 | 15.3 |
|  | 55+ | 462 | 6.43 |
|  | *missing = 3* | | |
| **Gender** | Male | 2433 | 33.94 |
|  | Female | 4726 | 66.06 |
|  | *missing = 16* |  |  |
| **Population group** | Black | 2428 | 33.83 |
|  | Coloured | 1257 | 17.51 |
|  | Indian | 823 | 11.47 |
|  | White | 2670 | 37.20 |
|  | *missing = 7* | | |
| **Marital status** | Single | 2625 | 36.6 |
|  | Married | 3823 | 53.3 |
|  | Divorced | 725 | 10.11 |
|  | *missing = 12* | | |
| **Number of dependents** | None | 1929 | 26.89 |
| 1-2 | 3694 | 51.50 |
|  | 3-4 | 1304 | 18.18 |
|  | 4+ | 246 | 3.43 |
|  | *missing=12* | | |
| **Income level** | < R5000 | 439 | 6.32 |
|  | R5001-R10000 | 2088 | 30.05 |
|  | R10001-R15000 | 1737 | 25 |
|  | R15001-R20000 | 1020 | 14.68 |
|  | R20001-R30000 | 925 | 13.31 |
|  | R30000+ | 740 | 10.65 |
|  | *missing = 236* | | |
| **Household earner** | Yes | 3268 | 45.52 |
|  | No | 3912 | 54.48 |
|  | *missing = 5* |  |  |

**Table 2: Frequency distribution and descriptive statistics for total SOC**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **SOC total\_rounded** | **Frequency** | **Percent** | **Cumulative Frequency** | **Cumulative Percent** |
| 1 | 150 | 2.32 | 150 | 2.32 |
| 2 | 368 | 5.69 | 518 | 8.01 |
| 3 | 733 | 11.34 | 1251 | 19.36 |
| 4 | 1299 | 20.10 | 2550 | 39.46 |
| 5 | 1602 | 24.79 | 4152 | 64.24 |
| 6 | 1577 | 24.40 | 5729 | 88.64 |
| 7 | 734 | 11.36 | 6463 | 100.00 |
| **Descriptive statistics** | | | | |
| **N** | 6463 | **Sum Weights** | | 6463 |
| **Mean** | 4.698 | **Sum Observations** | | 30366.333 |
| **Std Deviation** | 1.449 | **Variance** | | 2.099 |
| **Skewness** | -0.488 | **Kurtosis** | | -0.407 |
| **Uncorrected SS** | 156244.4 | **Corrected SS** | | 13568.522 |
| **Coeff Variation** | 30.840 | **Std Error Mean** | | 0.018 |

**Table 3: Summary of statistically significant predictor variables for high/low SOC**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **High/low SOC categories** | **Predictor variables entered** | **Predictor variables entered with statistical significance in each tree**  **level** | **tree levels** | **% correct high/low SOC classification** |
| SOCstddev | income, age, dependents, population group, marital status, gender + earner | I: income  II: marital ;earner  III: earner ;dependent | 3 | 69.1% |
| SOCMinus45 | income, age, dependents, population group, marital status, gender + earner | I: income  II: marital earner ;dependent; population  III: earner ;marital; dependent | 3 | 70.0% |