

Midterm Q1

Bhagyarathi Raman

3/7/2022

```
library(AER)
```

```
## Loading required package: car
```

```
## Loading required package: carData
```

```
## Loading required package: lmttest
```

```
## Loading required package: zoo
```

```
##  
## Attaching package: 'zoo'
```

```
## The following objects are masked from 'package:base':  
##  
## as.Date, as.Date.numeric
```

```
## Loading required package: sandwich
```

```
## Loading required package: survival
```

```
data(Affairs)  
names(Affairs)
```

```
## [1] "affairs"      "gender"      "age"      "yearsmarried"  
## [5] "children"    "religiousness" "education" "occupation"  
## [9] "rating"
```

```
head(Affairs)
```

```
##   affairs gender age yearsmarried children religiousness education occupation  
## 4      0  male 37      10.00      no      3      18      7  
## 5      0 female 27       4.00      no      4      14      6  
## 11     0 female 32      15.00     yes      1      12      1  
## 16     0  male 57      15.00     yes      5      18      6  
## 23     0  male 22       0.75      no      2      17      6  
## 29     0 female 32       1.50      no      2      17      5  
## rating  
## 4      4  
## 5      4  
## 11     4  
## 16     5  
## 23     3  
## 29     5
```

1. Process the data (5 points)

a. Missing values. How many variables have missing values?

```
is.na(Affairs)
```

```
##   affairs gender age yearsmarried children religiousness education  
## 4      FALSE FALSE FALSE      FALSE FALSE      FALSE      FALSE  
## 5      FALSE FALSE FALSE      FALSE FALSE      FALSE      FALSE  
## 11     FALSE FALSE FALSE      FALSE FALSE      FALSE      FALSE  
## 16     FALSE FALSE FALSE      FALSE FALSE      FALSE      FALSE  
## 23     FALSE FALSE FALSE      FALSE FALSE      FALSE      FALSE  
## 29     FALSE FALSE FALSE      FALSE FALSE      FALSE      FALSE  
## 44     FALSE FALSE FALSE      FALSE FALSE      FALSE      FALSE  
## 45     FALSE FALSE FALSE      FALSE FALSE      FALSE      FALSE  
## 47     FALSE FALSE FALSE      FALSE FALSE      FALSE      FALSE  
## 49     FALSE FALSE FALSE      FALSE FALSE      FALSE      FALSE
```

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

##	1564	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
##	1573	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
##	1575	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
##	1599	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
##	1622	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
##	1629	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
##	1664	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
##	1669	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
##	1674	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
##	1682	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
##	1685	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
##	1697	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
##	1716	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
##	1730	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
##	1731	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
##	1732	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
##	1743	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
##	1751	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
##	1757	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
##	1763	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
##	1766	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
##	1772	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
##	1776	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
##	1782	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
##	1784	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
##	1791	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
##	1831	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
##	1840	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
##	1844	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
##	1856	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
##	1876	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
##	1929	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
##	1935	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
##	1938	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
##	1941	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
##	1954	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
##	1959	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
##	9010	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
##	occupation rating							
##	4	FALSE	FALSE					
##	5	FALSE	FALSE					
##	11	FALSE	FALSE					
##	16	FALSE	FALSE					
##	23	FALSE	FALSE					
##	29	FALSE	FALSE					
##	44	FALSE	FALSE					
##	45	FALSE	FALSE					
##	47	FALSE	FALSE					
##	49	FALSE	FALSE					
##	50	FALSE	FALSE					
##	55	FALSE	FALSE					
##	64	FALSE	FALSE					
##	80	FALSE	FALSE					
##	86	FALSE	FALSE					
##	93	FALSE	FALSE					
##	108	FALSE	FALSE					
##	114	FALSE	FALSE					
##	115	FALSE	FALSE					
##	116	FALSE	FALSE					
##	123	FALSE	FALSE					
##	127	FALSE	FALSE					
##	129	FALSE	FALSE					
##	134	FALSE	FALSE					
##	137	FALSE	FALSE					
##	139	FALSE	FALSE					
##	147	FALSE	FALSE					
##	151	FALSE	FALSE					
##	153	FALSE	FALSE					
##	155	FALSE	FALSE					
##	162	FALSE	FALSE					
##	163	FALSE	FALSE					
##	165	FALSE	FALSE					
##	168	FALSE	FALSE					
##	170	FALSE	FALSE					
##	172	FALSE	FALSE					
##	184	FALSE	FALSE					
##	187	FALSE	FALSE					
##	192	FALSE	FALSE					
##	194	FALSE	FALSE					

## 210	FALSE	FALSE
## 217	FALSE	FALSE
## 220	FALSE	FALSE
## 224	FALSE	FALSE
## 227	FALSE	FALSE
## 228	FALSE	FALSE
## 239	FALSE	FALSE
## 241	FALSE	FALSE
## 245	FALSE	FALSE
## 249	FALSE	FALSE
## 262	FALSE	FALSE
## 265	FALSE	FALSE
## 267	FALSE	FALSE
## 269	FALSE	FALSE
## 271	FALSE	FALSE
## 277	FALSE	FALSE
## 290	FALSE	FALSE
## 292	FALSE	FALSE
## 293	FALSE	FALSE
## 295	FALSE	FALSE
## 299	FALSE	FALSE
## 320	FALSE	FALSE
## 321	FALSE	FALSE
## 324	FALSE	FALSE
## 334	FALSE	FALSE
## 351	FALSE	FALSE
## 355	FALSE	FALSE
## 361	FALSE	FALSE
## 362	FALSE	FALSE
## 366	FALSE	FALSE
## 370	FALSE	FALSE
## 374	FALSE	FALSE
## 378	FALSE	FALSE
## 381	FALSE	FALSE
## 382	FALSE	FALSE
## 383	FALSE	FALSE
## 384	FALSE	FALSE
## 400	FALSE	FALSE
## 403	FALSE	FALSE
## 409	FALSE	FALSE
## 412	FALSE	FALSE
## 413	FALSE	FALSE
## 416	FALSE	FALSE
## 418	FALSE	FALSE
## 422	FALSE	FALSE
## 435	FALSE	FALSE
## 439	FALSE	FALSE
## 445	FALSE	FALSE
## 447	FALSE	FALSE
## 448	FALSE	FALSE
## 449	FALSE	FALSE
## 478	FALSE	FALSE
## 482	FALSE	FALSE
## 486	FALSE	FALSE
## 489	FALSE	FALSE
## 490	FALSE	FALSE
## 491	FALSE	FALSE
## 492	FALSE	FALSE
## 503	FALSE	FALSE
## 508	FALSE	FALSE
## 509	FALSE	FALSE
## 512	FALSE	FALSE
## 515	FALSE	FALSE
## 517	FALSE	FALSE
## 532	FALSE	FALSE
## 533	FALSE	FALSE
## 535	FALSE	FALSE
## 537	FALSE	FALSE
## 538	FALSE	FALSE
## 543	FALSE	FALSE
## 547	FALSE	FALSE
## 550	FALSE	FALSE
## 558	FALSE	FALSE
## 571	FALSE	FALSE
## 578	FALSE	FALSE
## 583	FALSE	FALSE
## 586	FALSE	FALSE
## 594	FALSE	FALSE
## 597	FALSE	FALSE

## 602	FALSE	FALSE
## 603	FALSE	FALSE
## 604	FALSE	FALSE
## 612	FALSE	FALSE
## 613	FALSE	FALSE
## 621	FALSE	FALSE
## 627	FALSE	FALSE
## 630	FALSE	FALSE
## 631	FALSE	FALSE
## 632	FALSE	FALSE
## 639	FALSE	FALSE
## 645	FALSE	FALSE
## 647	FALSE	FALSE
## 648	FALSE	FALSE
## 651	FALSE	FALSE
## 655	FALSE	FALSE
## 667	FALSE	FALSE
## 670	FALSE	FALSE
## 671	FALSE	FALSE
## 673	FALSE	FALSE
## 701	FALSE	FALSE
## 705	FALSE	FALSE
## 706	FALSE	FALSE
## 709	FALSE	FALSE
## 717	FALSE	FALSE
## 719	FALSE	FALSE
## 723	FALSE	FALSE
## 724	FALSE	FALSE
## 726	FALSE	FALSE
## 734	FALSE	FALSE
## 735	FALSE	FALSE
## 736	FALSE	FALSE
## 737	FALSE	FALSE
## 739	FALSE	FALSE
## 743	FALSE	FALSE
## 745	FALSE	FALSE
## 747	FALSE	FALSE
## 751	FALSE	FALSE
## 752	FALSE	FALSE
## 754	FALSE	FALSE
## 760	FALSE	FALSE
## 763	FALSE	FALSE
## 774	FALSE	FALSE
## 776	FALSE	FALSE
## 779	FALSE	FALSE
## 784	FALSE	FALSE
## 788	FALSE	FALSE
## 794	FALSE	FALSE
## 795	FALSE	FALSE
## 798	FALSE	FALSE
## 800	FALSE	FALSE
## 803	FALSE	FALSE
## 807	FALSE	FALSE
## 812	FALSE	FALSE
## 820	FALSE	FALSE
## 823	FALSE	FALSE
## 830	FALSE	FALSE
## 843	FALSE	FALSE
## 848	FALSE	FALSE
## 851	FALSE	FALSE
## 854	FALSE	FALSE
## 856	FALSE	FALSE
## 857	FALSE	FALSE
## 859	FALSE	FALSE
## 863	FALSE	FALSE
## 865	FALSE	FALSE
## 867	FALSE	FALSE
## 870	FALSE	FALSE
## 873	FALSE	FALSE
## 875	FALSE	FALSE
## 876	FALSE	FALSE
## 877	FALSE	FALSE
## 880	FALSE	FALSE
## 903	FALSE	FALSE
## 904	FALSE	FALSE
## 905	FALSE	FALSE
## 908	FALSE	FALSE
## 909	FALSE	FALSE
## 910	FALSE	FALSE

## 912	FALSE	FALSE
## 914	FALSE	FALSE
## 915	FALSE	FALSE
## 916	FALSE	FALSE
## 920	FALSE	FALSE
## 921	FALSE	FALSE
## 925	FALSE	FALSE
## 926	FALSE	FALSE
## 929	FALSE	FALSE
## 931	FALSE	FALSE
## 945	FALSE	FALSE
## 947	FALSE	FALSE
## 949	FALSE	FALSE
## 950	FALSE	FALSE
## 961	FALSE	FALSE
## 965	FALSE	FALSE
## 966	FALSE	FALSE
## 967	FALSE	FALSE
## 987	FALSE	FALSE
## 990	FALSE	FALSE
## 992	FALSE	FALSE
## 995	FALSE	FALSE
## 1009	FALSE	FALSE
## 1021	FALSE	FALSE
## 1026	FALSE	FALSE
## 1027	FALSE	FALSE
## 1030	FALSE	FALSE
## 1031	FALSE	FALSE
## 1034	FALSE	FALSE
## 1037	FALSE	FALSE
## 1038	FALSE	FALSE
## 1039	FALSE	FALSE
## 1045	FALSE	FALSE
## 1046	FALSE	FALSE
## 1054	FALSE	FALSE
## 1059	FALSE	FALSE
## 1063	FALSE	FALSE
## 1068	FALSE	FALSE
## 1070	FALSE	FALSE
## 1072	FALSE	FALSE
## 1073	FALSE	FALSE
## 1077	FALSE	FALSE
## 1081	FALSE	FALSE
## 1083	FALSE	FALSE
## 1084	FALSE	FALSE
## 1086	FALSE	FALSE
## 1087	FALSE	FALSE
## 1089	FALSE	FALSE
## 1096	FALSE	FALSE
## 1102	FALSE	FALSE
## 1103	FALSE	FALSE
## 1107	FALSE	FALSE
## 1109	FALSE	FALSE
## 1115	FALSE	FALSE
## 1119	FALSE	FALSE
## 1124	FALSE	FALSE
## 1126	FALSE	FALSE
## 1128	FALSE	FALSE
## 1129	FALSE	FALSE
## 1130	FALSE	FALSE
## 1133	FALSE	FALSE
## 1140	FALSE	FALSE
## 1143	FALSE	FALSE
## 1146	FALSE	FALSE
## 1153	FALSE	FALSE
## 1156	FALSE	FALSE
## 1157	FALSE	FALSE
## 1158	FALSE	FALSE
## 1160	FALSE	FALSE
## 1161	FALSE	FALSE
## 1166	FALSE	FALSE
## 1177	FALSE	FALSE
## 1178	FALSE	FALSE
## 1180	FALSE	FALSE
## 1187	FALSE	FALSE
## 1191	FALSE	FALSE
## 1195	FALSE	FALSE
## 1207	FALSE	FALSE
## 1208	FALSE	FALSE

##	1209	FALSE	FALSE
##	1211	FALSE	FALSE
##	1215	FALSE	FALSE
##	1221	FALSE	FALSE
##	1226	FALSE	FALSE
##	1229	FALSE	FALSE
##	1231	FALSE	FALSE
##	1234	FALSE	FALSE
##	1235	FALSE	FALSE
##	1242	FALSE	FALSE
##	1245	FALSE	FALSE
##	1260	FALSE	FALSE
##	1266	FALSE	FALSE
##	1271	FALSE	FALSE
##	1273	FALSE	FALSE
##	1276	FALSE	FALSE
##	1280	FALSE	FALSE
##	1282	FALSE	FALSE
##	1285	FALSE	FALSE
##	1295	FALSE	FALSE
##	1298	FALSE	FALSE
##	1299	FALSE	FALSE
##	1304	FALSE	FALSE
##	1305	FALSE	FALSE
##	1311	FALSE	FALSE
##	1314	FALSE	FALSE
##	1319	FALSE	FALSE
##	1322	FALSE	FALSE
##	1324	FALSE	FALSE
##	1327	FALSE	FALSE
##	1328	FALSE	FALSE
##	1330	FALSE	FALSE
##	1332	FALSE	FALSE
##	1333	FALSE	FALSE
##	1336	FALSE	FALSE
##	1341	FALSE	FALSE
##	1344	FALSE	FALSE
##	1352	FALSE	FALSE
##	1358	FALSE	FALSE
##	1359	FALSE	FALSE
##	1361	FALSE	FALSE
##	1364	FALSE	FALSE
##	1368	FALSE	FALSE
##	1384	FALSE	FALSE
##	1390	FALSE	FALSE
##	1393	FALSE	FALSE
##	1394	FALSE	FALSE
##	1402	FALSE	FALSE
##	1407	FALSE	FALSE
##	1408	FALSE	FALSE
##	1412	FALSE	FALSE
##	1413	FALSE	FALSE
##	1416	FALSE	FALSE
##	1417	FALSE	FALSE
##	1418	FALSE	FALSE
##	1419	FALSE	FALSE
##	1420	FALSE	FALSE
##	1423	FALSE	FALSE
##	1424	FALSE	FALSE
##	1432	FALSE	FALSE
##	1433	FALSE	FALSE
##	1437	FALSE	FALSE
##	1438	FALSE	FALSE
##	1439	FALSE	FALSE
##	1446	FALSE	FALSE
##	1450	FALSE	FALSE
##	1451	FALSE	FALSE
##	1452	FALSE	FALSE
##	1453	FALSE	FALSE
##	1456	FALSE	FALSE
##	1464	FALSE	FALSE
##	1469	FALSE	FALSE
##	1473	FALSE	FALSE
##	1481	FALSE	FALSE
##	1482	FALSE	FALSE
##	1496	FALSE	FALSE
##	1497	FALSE	FALSE
##	1504	FALSE	FALSE
##	1513	FALSE	FALSE

##	1515	FALSE	FALSE
##	1534	FALSE	FALSE
##	1535	FALSE	FALSE
##	1536	FALSE	FALSE
##	1540	FALSE	FALSE
##	1551	FALSE	FALSE
##	1555	FALSE	FALSE
##	1557	FALSE	FALSE
##	1566	FALSE	FALSE
##	1567	FALSE	FALSE
##	1576	FALSE	FALSE
##	1584	FALSE	FALSE
##	1585	FALSE	FALSE
##	1590	FALSE	FALSE
##	1594	FALSE	FALSE
##	1595	FALSE	FALSE
##	1603	FALSE	FALSE
##	1608	FALSE	FALSE
##	1609	FALSE	FALSE
##	1615	FALSE	FALSE
##	1616	FALSE	FALSE
##	1617	FALSE	FALSE
##	1620	FALSE	FALSE
##	1621	FALSE	FALSE
##	1637	FALSE	FALSE
##	1638	FALSE	FALSE
##	1650	FALSE	FALSE
##	1654	FALSE	FALSE
##	1665	FALSE	FALSE
##	1670	FALSE	FALSE
##	1671	FALSE	FALSE
##	1675	FALSE	FALSE
##	1688	FALSE	FALSE
##	1691	FALSE	FALSE
##	1695	FALSE	FALSE
##	1698	FALSE	FALSE
##	1704	FALSE	FALSE
##	1705	FALSE	FALSE
##	1711	FALSE	FALSE
##	1719	FALSE	FALSE
##	1723	FALSE	FALSE
##	1726	FALSE	FALSE
##	1749	FALSE	FALSE
##	1752	FALSE	FALSE
##	1754	FALSE	FALSE
##	1758	FALSE	FALSE
##	1761	FALSE	FALSE
##	1773	FALSE	FALSE
##	1775	FALSE	FALSE
##	1786	FALSE	FALSE
##	1793	FALSE	FALSE
##	1799	FALSE	FALSE
##	1803	FALSE	FALSE
##	1806	FALSE	FALSE
##	1807	FALSE	FALSE
##	1808	FALSE	FALSE
##	1814	FALSE	FALSE
##	1815	FALSE	FALSE
##	1818	FALSE	FALSE
##	1827	FALSE	FALSE
##	1834	FALSE	FALSE
##	1835	FALSE	FALSE
##	1843	FALSE	FALSE
##	1846	FALSE	FALSE
##	1850	FALSE	FALSE
##	1851	FALSE	FALSE
##	1854	FALSE	FALSE
##	1859	FALSE	FALSE
##	1861	FALSE	FALSE
##	1866	FALSE	FALSE
##	1873	FALSE	FALSE
##	1875	FALSE	FALSE
##	1885	FALSE	FALSE
##	1892	FALSE	FALSE
##	1895	FALSE	FALSE
##	1896	FALSE	FALSE
##	1897	FALSE	FALSE
##	1899	FALSE	FALSE
##	1904	FALSE	FALSE

## 1905	FALSE	FALSE
## 1908	FALSE	FALSE
## 1916	FALSE	FALSE
## 1918	FALSE	FALSE
## 1920	FALSE	FALSE
## 1930	FALSE	FALSE
## 1940	FALSE	FALSE
## 1947	FALSE	FALSE
## 1949	FALSE	FALSE
## 1951	FALSE	FALSE
## 1952	FALSE	FALSE
## 1960	FALSE	FALSE
## 9001	FALSE	FALSE
## 9012	FALSE	FALSE
## 9023	FALSE	FALSE
## 9029	FALSE	FALSE
## 6	FALSE	FALSE
## 12	FALSE	FALSE
## 43	FALSE	FALSE
## 53	FALSE	FALSE
## 67	FALSE	FALSE
## 79	FALSE	FALSE
## 122	FALSE	FALSE
## 126	FALSE	FALSE
## 133	FALSE	FALSE
## 138	FALSE	FALSE
## 154	FALSE	FALSE
## 159	FALSE	FALSE
## 174	FALSE	FALSE
## 176	FALSE	FALSE
## 181	FALSE	FALSE
## 182	FALSE	FALSE
## 186	FALSE	FALSE
## 189	FALSE	FALSE
## 204	FALSE	FALSE
## 215	FALSE	FALSE
## 232	FALSE	FALSE
## 233	FALSE	FALSE
## 252	FALSE	FALSE
## 253	FALSE	FALSE
## 274	FALSE	FALSE
## 275	FALSE	FALSE
## 287	FALSE	FALSE
## 288	FALSE	FALSE
## 325	FALSE	FALSE
## 328	FALSE	FALSE
## 344	FALSE	FALSE
## 353	FALSE	FALSE
## 354	FALSE	FALSE
## 367	FALSE	FALSE
## 369	FALSE	FALSE
## 390	FALSE	FALSE
## 392	FALSE	FALSE
## 423	FALSE	FALSE
## 432	FALSE	FALSE
## 436	FALSE	FALSE
## 483	FALSE	FALSE
## 513	FALSE	FALSE
## 516	FALSE	FALSE
## 518	FALSE	FALSE
## 520	FALSE	FALSE
## 526	FALSE	FALSE
## 528	FALSE	FALSE
## 553	FALSE	FALSE
## 576	FALSE	FALSE
## 611	FALSE	FALSE
## 625	FALSE	FALSE
## 635	FALSE	FALSE
## 646	FALSE	FALSE
## 657	FALSE	FALSE
## 659	FALSE	FALSE
## 666	FALSE	FALSE
## 679	FALSE	FALSE
## 729	FALSE	FALSE
## 755	FALSE	FALSE
## 758	FALSE	FALSE
## 770	FALSE	FALSE
## 786	FALSE	FALSE
## 797	FALSE	FALSE

## 811	FALSE	FALSE
## 834	FALSE	FALSE
## 858	FALSE	FALSE
## 885	FALSE	FALSE
## 893	FALSE	FALSE
## 927	FALSE	FALSE
## 928	FALSE	FALSE
## 933	FALSE	FALSE
## 951	FALSE	FALSE
## 968	FALSE	FALSE
## 972	FALSE	FALSE
## 975	FALSE	FALSE
## 977	FALSE	FALSE
## 981	FALSE	FALSE
## 986	FALSE	FALSE
## 1002	FALSE	FALSE
## 1007	FALSE	FALSE
## 1011	FALSE	FALSE
## 1035	FALSE	FALSE
## 1050	FALSE	FALSE
## 1056	FALSE	FALSE
## 1057	FALSE	FALSE
## 1075	FALSE	FALSE
## 1080	FALSE	FALSE
## 1125	FALSE	FALSE
## 1131	FALSE	FALSE
## 1138	FALSE	FALSE
## 1150	FALSE	FALSE
## 1163	FALSE	FALSE
## 1169	FALSE	FALSE
## 1198	FALSE	FALSE
## 1204	FALSE	FALSE
## 1218	FALSE	FALSE
## 1230	FALSE	FALSE
## 1236	FALSE	FALSE
## 1247	FALSE	FALSE
## 1259	FALSE	FALSE
## 1294	FALSE	FALSE
## 1353	FALSE	FALSE
## 1370	FALSE	FALSE
## 1427	FALSE	FALSE
## 1445	FALSE	FALSE
## 1460	FALSE	FALSE
## 1480	FALSE	FALSE
## 1505	FALSE	FALSE
## 1543	FALSE	FALSE
## 1548	FALSE	FALSE
## 1550	FALSE	FALSE
## 1561	FALSE	FALSE
## 1564	FALSE	FALSE
## 1573	FALSE	FALSE
## 1575	FALSE	FALSE
## 1599	FALSE	FALSE
## 1622	FALSE	FALSE
## 1629	FALSE	FALSE
## 1664	FALSE	FALSE
## 1669	FALSE	FALSE
## 1674	FALSE	FALSE
## 1682	FALSE	FALSE
## 1685	FALSE	FALSE
## 1697	FALSE	FALSE
## 1716	FALSE	FALSE
## 1730	FALSE	FALSE
## 1731	FALSE	FALSE
## 1732	FALSE	FALSE
## 1743	FALSE	FALSE
## 1751	FALSE	FALSE
## 1757	FALSE	FALSE
## 1763	FALSE	FALSE
## 1766	FALSE	FALSE
## 1772	FALSE	FALSE
## 1776	FALSE	FALSE
## 1782	FALSE	FALSE
## 1784	FALSE	FALSE
## 1791	FALSE	FALSE
## 1831	FALSE	FALSE
## 1840	FALSE	FALSE
## 1844	FALSE	FALSE
## 1856	FALSE	FALSE


```
## 1876      FALSE FALSE
## 1929      FALSE FALSE
## 1935      FALSE FALSE
## 1938      FALSE FALSE
## 1941      FALSE FALSE
## 1954      FALSE FALSE
## 1959      FALSE FALSE
## 9010      FALSE FALSE
```

```
#There are no variables having missing values in this data.
```

b. If the data have missing values, explain what might be a better approach than simply deleting data with missing values.

```
#The data does not have any missing values. If there is a missing values in any data, the better approach is to use "NA" or "NAN" rather than deleting the values.
```

c. Factors: Which variables are factors? Demonstrate your answer using the function str().

```
str(Affairs)
```

```
## 'data.frame':   601 obs. of  9 variables:
## $ affairs      : num  0 0 0 0 0 0 0 0 0 ...
## $ gender       : Factor w/ 2 levels "female","male": 2 1 1 2 2 1 1 2 1 2 ...
## $ age          : num  37 27 32 57 22 32 22 57 32 22 ...
## $ yearsmarried : num  10 4 15 15 0.75 1.5 0.75 15 15 1.5 ...
## $ children     : Factor w/ 2 levels "no","yes": 1 1 2 2 1 1 1 2 2 1 ...
## $ religiousness: int   3 4 1 5 2 2 2 2 4 4 ...
## $ education    : num  18 14 12 18 17 17 12 14 16 14 ...
## $ occupation   : int   7 6 1 6 6 5 1 4 1 4 ...
## $ rating       : int   4 4 4 5 3 5 3 4 2 5 ...
```

```
#Gender and Children variable are factors.
```

d. Obtain a summary of the data. Does there appear to be any anomalies?

```
summary(Affairs)
```

```
##      affairs      gender      age      yearsmarried      children
## Min.   : 0.000  female:315  Min.   :17.50  Min.   : 0.125  no :171
## 1st Qu.: 0.000  male  :286  1st Qu.:27.00  1st Qu.: 4.000  yes:430
## Median : 0.000                Median :32.00  Median : 7.000
## Mean   : 1.456                Mean   :32.49  Mean   : 8.178
## 3rd Qu.: 0.000                3rd Qu.:37.00  3rd Qu.:15.000
## Max.   :12.000                Max.   :57.00  Max.   :15.000
## religiousness  education      occupation      rating
## Min.   :1.000  Min.   : 9.00  Min.   :1.000  Min.   :1.000
## 1st Qu.:2.000  1st Qu.:14.00  1st Qu.:3.000  1st Qu.:3.000
## Median :3.000  Median :16.00  Median :5.000  Median :4.000
## Mean   :3.116  Mean   :16.17  Mean   :4.195  Mean   :3.932
## 3rd Qu.:4.000  3rd Qu.:18.00  3rd Qu.:6.000  3rd Qu.:5.000
## Max.   :5.000  Max.   :20.00  Max.   :7.000  Max.   :5.000
```

```
# No anomalies are seem to be appeared.
```

e. Obtain the correlation matrix and correlation plot of the numeric variables. Which predictor has the largest, positive or negative, correlation with affairs?

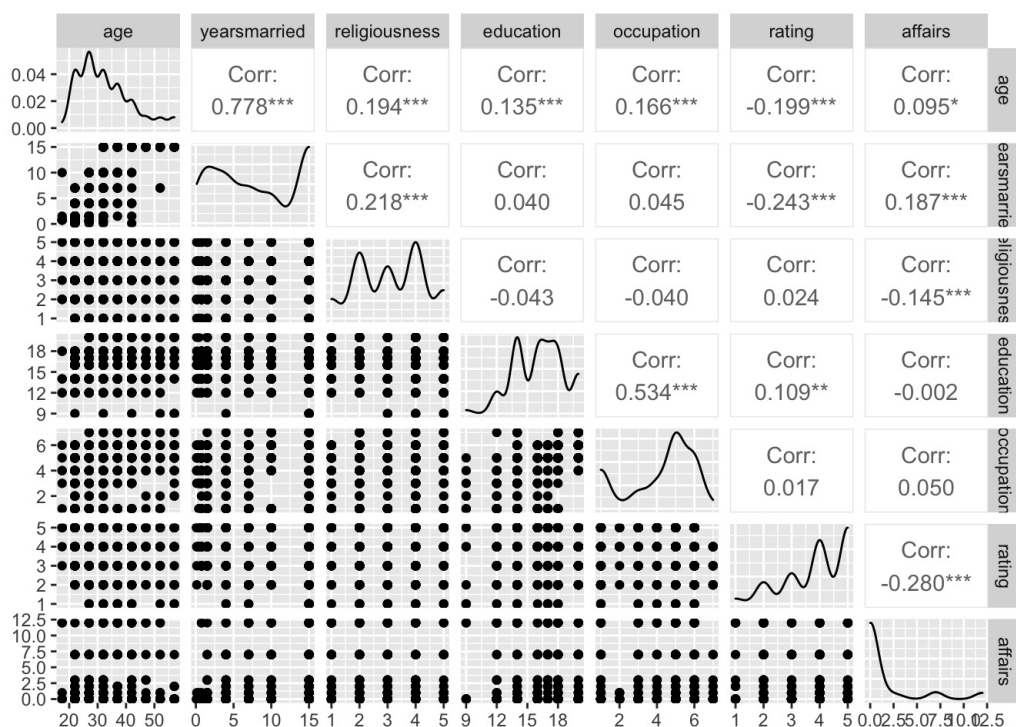
```
library(ggplot2)
cor(Affairs[sapply(Affairs,is.numeric)])
```

```
##          affairs      age yearsmarried religiousness  education
## affairs      1.00000000  0.0952372   0.18684169  -0.14450135 -0.002437441
## age          0.095237204  1.0000000   0.77754585   0.19377693  0.134596015
## yearsmarried 0.186841686  0.7775458   1.00000000   0.21826067  0.040002716
## religiousness -0.144501345  0.1937769  0.21826067   1.00000000 -0.042571079
## education    -0.002437441  0.1345960  0.04000272  -0.04257108  1.000000000
## occupation    0.049611758  0.1664125  0.04459201  -0.03972232  0.533605242
## rating        -0.279512403 -0.1989999  -0.24311883   0.02429578  0.109303473
##
##          occupation      rating
## affairs      0.04961176 -0.27951240
## age          0.16641254 -0.19899990
## yearsmarried 0.04459201 -0.24311883
## religiousness -0.03972232  0.02429578
## education    0.53360524  0.10930347
## occupation    1.00000000  0.01742227
## rating        0.01742227  1.00000000
```

```
library(GGally)
```

```
## Registered S3 method overwritten by 'GGally':
##   method from
##   +.gg      ggplot2
```

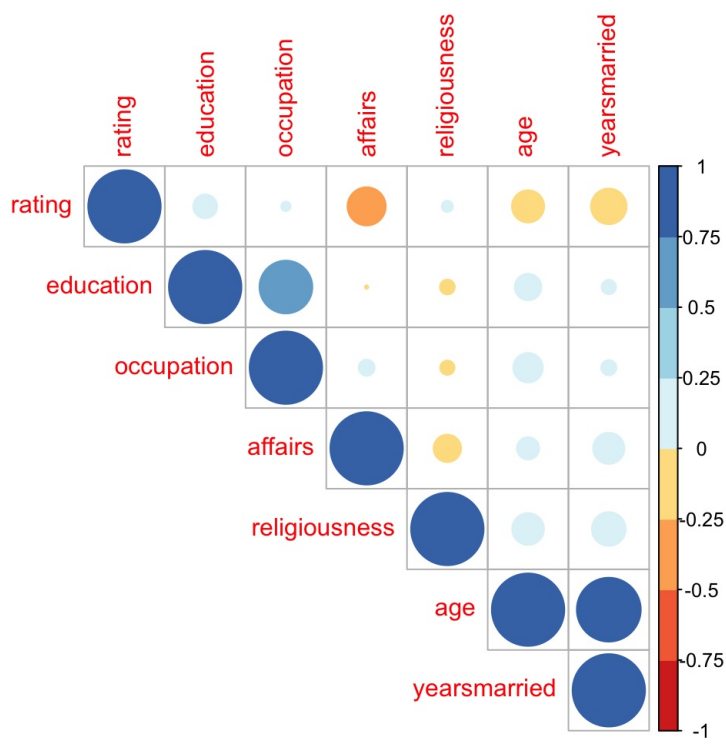
```
ggpairs(Affairs, columns=c("age",
                           "yearsmarried",
                           "religiousness",
                           "education",
                           "occupation",
                           "rating",
                           "affairs"))
```



```
library(corrplot)
```

```
## corrplot 0.92 loaded
```

```
library(RColorBrewer)
A <- cor(Affairs[sapply(Affairs, is.numeric)])
corrplot(A, type="upper", order="hclust",
         col=brewer.pal(n=8, name="RdYlBu"))
```



#Yearsmarried has the largest positive correlation with affairs and there is not a real negative correlation with affairs, but least correlation with affairs is rating.