

Sex Income

M 40.6

M 54.6

M 38.6

M 58.2

M 34.6

M 42.9

M 67.5

M 79.8

M 54.4

M 47.3

M 66.4

M 69.0

M 62.0

M 52.5

M 72.6

M 52.4

M 59.5

M 59.1

M 36.7

M 54.6

M 52.1

M 49.9

M 52.0

M 47.1

M 40.8

M 36.5

M 57.1

M 54.1

M 32.4

M 34.9

M 64.1

M 54.0

M 51.5

M 50.8

M 45.1

M 81.5

M 70.4

M 39.2

M 45.2

M 80.9

M 48.6

M 31.0

M 32.1

M 33.9

M 31.3

M 51.0

M 53.4

M 58.3

M 31.4

M 58.3

M 41.0

M 47.9

M 51.4

M 33.1

M 74.9

M 77.2

M 57.9

M 80.1

M 40.2

M 100.9

F 33.1

F 35.8

F 68.8

F 31.6

F 38.2

F 42.0

F 33.4

F 50.3

F 39.6

F 30.7

F 31.3

F 61.3

F 30.0

F 38.1

F 56.4

F 35.7

F 31.3

F 40.4

F 32.1

F 66.4

F 36.9

F 35.9

F 49.6

F 62.8

F 44.6

F 32.5

F 33.4

F 55.3

F 62.7

F 54.4

F 30.8

F 49.1

F 41.9

F 32.5

F 35.2

F 47.4

F 60.7

F 33.0

F 43.3

F 34.8

F 36.0

F 51.6

F 31.9

F 34.1

F 78.4

F 30.4

F 45.3

F 52.6

F 30.3

F 36.6

F 53.1

F 36.5

F 37.8

F 34.0

F 69.3

F 77.2

F 32.6

F 82.9

F 42.3

F 57.8

Variable 1

Male

n 60

Mean 52.913

SD 15.269

Variable 2

Female

n 60

Mean 44.233

SD 13.790

F-Test Two-Sample for Variances

	Variable 1	Variable 2
Mean	52.91333333	44.23333333
Variance	233.1289718	190.1758192
Observations	60	60
df	59	59
F	1.225860221	
P(F<=f) one-tail	0.21824624	
F Critical one-tail	1.539956607	

p2 0.43649248

t-Test: Two-Sample Assuming Equal Variances

	Variable 1	Variable 2
Mean	52.91333333	44.23333333
Variance	233.1289718	190.1758192
Observations	60	60
Pooled Variance	211.6523955	
Hypothesized Mean Difference	0	
df	118	
t Stat	3.267900001	
P(T<=t) one-tail	0.000709735	
t Critical one-tail	1.657869522	
P(T<=t) two-tail	0.00141947	
t Critical two-tail	1.980272249	

Difference in Means 8.68

Interpretation

A two-sample t-test (assuming equal variances) indicates that, on average, males earn about 8.68 more than females, with the test statistic providing strong evidence that this gap is unlikely to be due to chance ($p < 0.05$). In other words, the data support the claim that the population mean income for males exceeds that of females.

To justify the use of the equal-variances version of the t-test, the F-test for two-sample variances was performed. This procedure compares the variability of incomes in the male and female samples. Because the F-test shows no significant difference in how spread out the two distributions are, it is appropriate to combine the two variances into one pooled estimate.

The validity of this analysis relies on a few key assumptions. Each sample should be independently drawn from its respective population, and the data should be suitable for a t-test (meaning the sample sizes are large enough or the underlying distributions are not heavily skewed). By confirming these conditions and using the F-test results, we can be more confident that the observed difference in average incomes is genuine and not simply due to random variation.

DATA SET C (Superplus.xlsx)

This data set shows the incomes of 60 male and 60 female cardholders of a certain bank's "Superplus Diamond" card.

Variable **Description**
Sex The cardholder's sex (F = female, M = male)
Income The cardholder's gross annual income (in £'000's)

The data are as follows:

Sex	Income
M	40.6
M	54.6
M	38.6
?	?
F	33.1
F	35.8
F	68.8
?	?