Q1.

```
clc, clearvars, close all

weights_pounds = [122, 173, 179, 176, 159, 175, 160, 102, 133
159, 176, 151, 115, 105, 72, 170, 128, 112
101, 123, 117, 93, 117, 99, 90, 113, 128
129, 134, 178, 105, 107, 147, 157, 155, 95
177, 98, 174, 135, 97, 168, 160, 144, 174];

sum_wts_pds = 0;

for i = 1:5
    for j = 1:9
        sum_wts_pds = (sum_wts_pds + weights_pounds(i, j));
    end
end

mean_weight = (sum_wts_pds/45);
fprintf('Mean of the data is: %.4f\n', mean_weight);
```

Q2.

```
clc, clearvars, close all

f = [15, 20, 30, 18, 12, 5];
f_total = 0;
f_m_total = 0;

up_bd = [23, 28, 33, 38, 43, 48];
lw_bd = [19, 24, 29, 34, 39, 44];
midpts = ((up_bd + lw_bd)/2);
f_m = f .* midpts;

for i = 1:6
    f_total = f_total + f(i);
    f_m_total = f_m_total + f_m(i);
end

mean_value = (f_m_total/f_total);
fprintf('Mean of the data is: %.4f\n', mean_value);
```

Q3.

```
clc, clearvars, close all

weights_pounds = [122, 173, 179, 176, 159, 175, 160, 102, 133
159, 176, 151, 115, 105, 72, 170, 128, 112
101, 123, 117, 93, 117, 99, 90, 113, 128
129, 134, 178, 105, 107, 147, 157, 155, 95
177, 98, 174, 135, 97, 168, 160, 144, 174];

sum_wts_pds = 0;

for i = 1:5
    for j = 1:9
        sum_wts_pds = (sum_wts_pds + weights_pounds(i, j));
    end
end

mean_weight = (sum_wts_pds/45);

sum_dist_mean = 0;

for x = 1:5
    for y = 1:9
        dist_mean = ((weights_pounds(x, y) - mean_weight)^2);
        sum_dist_mean = (sum_dist_mean + dist_mean);
    end
end

variance = (sum_dist_mean/45);
fprintf('Variance of the data is: %.4f\n', variance);
```

Q4.

```
clc, clearvars, close all

f = [15, 20, 30, 18, 12, 5];
f_total = 0;
f_m_total = 0;

up_bd = [23, 28, 33, 38, 43, 48];
lw_bd = [19, 24, 29, 34, 39, 44];
midpts = ((up_bd + lw_bd)/2);
f_m = f .* midpts;

for i = 1:6
    f_total = f_total + f(i);
    f_m_total = f_m_total + f_m(i);
end

mean_value = (f_m_total/f_total);

midpt_x = zeros(1, 6);

for j = 1:6
    m_sub_x = (midpts(j) - mean_value);
    midpt_x(j) = m_sub_x;
end

midpt_x_sq = midpt_x .* midpt_x;
f_m_x = zeros(1, 6);

for k = 1:6
    f_m_x_val = f(k)*midpt_x_sq(k);
    f_m_x(k) = f_m_x_val;
end

f_m_x_total = 0;

for z = 1:6
    f_m_x_total = f_m_x_total + f_m_x(z);
end

variance = (f_m_x_total/f_total);
fprintf('Variance of the data is: %.4f\n', variance);
```

Q5.

```
clc, clearvars, close all

n_women = 12;
x_ages = [56, 42, 72, 36, 63, 47, 55, 49, 38, 42, 68, 60];
y_BP = [147, 125, 160, 118, 149, 128, 150, 145, 115, 140, 152, 155];

x_y = zeros(1, 12);
x_ages_sum = 0;
y_BP_sum = 0;
xy_sum = 0;

for i = 1:12
    x_ages_sum = x_ages_sum + x_ages(i);
    y_BP_sum = y_BP_sum + y_BP(i);
    x_y(i) = (x_ages(i) * y_BP(i));
end

for j = 1:12
    xy_sum = xy_sum + x_y(j);
end

x_mean = x_ages_sum / n_women;
y_mean = y_BP_sum / n_women;

covariance = (xy_sum - n_women * x_mean * y_mean) / n_women;

fprintf('Covariance between x and y: %.4f\n', covariance);
```

Q6.

```
clc, clearvars, close all

n = 100;
f_vals = [0 0 0 2 4 4
0 0 1 4 6 5
0 0 5 10 8 1
1 4 9 5 2 0
3 6 6 2 0 0
3 5 4 0 0 0];

% initializing upper and lower bounds for x & y
up_bd_x = [49 59 69 79 89 99];
up_bd_y = [99 89 79 69 59 49];
lw_bd_x = [40 50 60 70 80 90];
lw_bd_y = [90 80 70 60 50 40];

% initialiazing midpts and assumed means (same for x & y)
midpts_x = (up_bd_x + lw_bd_x) / 2;
midpts_y = (up_bd_y + lw_bd_y) / 2;
mean_asmd_xy = 74.5;
% obtaining factor from subtraction from assumed mean
diff_factor = 10;

% calculating u & v vals
u_vals = zeros(1, 6);

for i = 1:6
    u_vals(i) = ((midpts_x(i) - mean_asmd_xy) / diff_factor);
end

v_vals = u_vals(end:-1:1);

% initializing a 6x6 zero matrix to store fuv values
fuv = zeros(6, 6);
% following loop traverses through each cell and stores the required
value
for j = 1:6
    for k = 1:6
        fuv(j, k) = f_vals(j, k) * v_vals(j) * u_vals(k);
    end
end



% initializing required matrices
% for v
f_v = [10 16 24 21 17 12]; % given
```

```
fv = zeros(1, 6);
fv_sq = zeros(1, 6);
f_uv = 0;

%for u
f_u = [7 15 25 23 20 10]; % given
fu = zeros(1, 6);
fu_sq = zeros(1, 6);

% calculating fu, fv, fu^2, fv^2
for a = 1:6
    fv(a) = f_v(a) .* v_vals(a);
    fv_sq(a) = f_v(a) .* (v_vals(a)^2);
    fu(a) = f_u(a) .* u_vals(a);
    fu_sq(a) = f_u(a) .* (u_vals(a)^2);
end

% calculating f_uv
for m = 1:36
    f_uv = f_uv + fuv(m);
end
% calculating respective sums
fv_sum = 0;
fv_sq_sum = 0;
fu_sum = 0;
fu_sq_sum = 0;

for d = 1:6
    fv_sum =fv_sum + fv(d);
    fv_sq_sum =fv_sq_sum + fv_sq(d);
    fu_sum =fu_sum + fu(d);
    fu_sq_sum =fu_sq_sum + fu_sq(d);
end

% calculating square of summation of fu & fv respectively
fu_sum_sq = fu_sum ^ 2;
fv_sum_sq = fv_sum ^ 2;

% calculating co-variance between x & y
corr_coeff_num = (n*f_uv) - (fu_sum*fv_sum);
covar_xy = corr_coeff_num/100;
% corr_coeff_denom = (sqrt((n*fu_sq_sum)-
fu_sum_sq))*(sqrt((n*fv_sq_sum)-fv_sum_sq));
% corr_coeff = corr_coeff_num / corr_coeff_denom;

fprintf('Co-variance between the marks in Mathematics and the marks in
Physics is: %.4f\n', covar_xy);
```

Q7.

```
clc, clearvars, close all

n_women = 12;
x_ages = [56, 42, 72, 36, 63, 47, 55, 49, 38, 42, 68, 60];
y_BP = [147, 125, 160, 118, 149, 128, 150, 145, 115, 140, 152, 155];

x_y = zeros(1, 12);
x_sq = zeros(1, 12);
y_sq = zeros(1, 12);

for i = 1:12
    x_y(i) = (x_ages(i) * y_BP(i));
    x_sq(i) = (x_ages(i) ^ 2);
    y_sq(i) = (y_BP(i) ^ 2);
end

x_ages_sum = 0;
y_BP_sum = 0;
x_y_sum = 0;
x_sq_sum = 0;
y_sq_sum = 0;

for j = 1:12
    x_ages_sum = x_ages_sum + x_ages(j);
    y_BP_sum = y_BP_sum + y_BP(j);
    x_y_sum = x_y_sum + x_y(j);
    x_sq_sum = x_sq_sum + x_sq(j);
    y_sq_sum = y_sq_sum + y_sq(j);
end

r_numerator = (n_women*(x_y_sum)) - (x_ages_sum * y_BP_sum);
r_denominator_1 = (n_women*x_sq_sum) - (x_ages_sum^2);
r_denominator_2 = (n_women*y_sq_sum) - (y_BP_sum^2);
r = r_denominator_1*r_denominator_2;
r = r^0.5;
r = r_numerator / r;

fprintf('Correlation coefficient between x and y: %.4f\n', r);
```

Q8.

```
clc, clearvars, close all

n = 100;
f_vals = [0 0 0 2 4 4
0 0 1 4 6 5
0 0 5 10 8 1
1 4 9 5 2 0
3 6 6 2 0 0
3 5 4 0 0 0];

% initializing upper and lower bounds for x & y
up_bd_x = [49 59 69 79 89 99];
up_bd_y = [99 89 79 69 59 49];
lw_bd_x = [40 50 60 70 80 90];
lw_bd_y = [90 80 70 60 50 40];

% initialiazing midpts and assumed means (same for x & y)
midpts_x = (up_bd_x + lw_bd_x) / 2;
midpts_y = (up_bd_y + lw_bd_y) / 2;
mean_asmd_xy = 74.5;
% obtaining factor from subtraction from assumed mean
diff_factor = 10;

% calculating u & v vals
u_vals = zeros(1, 6);

for i = 1:6
    u_vals(i) = ((midpts_x(i) - mean_asmd_xy) / diff_factor);
end

v_vals = u_vals(end:-1:1);

% initializing a 6x6 zero matrix to store fuv values
fuv = zeros(6, 6);
% following loop traverses through each cell and stores the required
value
for j = 1:6
    for k = 1:6
        fuv(j, k) = f_vals(j, k) * v_vals(j) * u_vals(k);
    end
end



% initializing required matrices
% for v
f_v = [10 16 24 21 17 12]; % given
```

```
fv = zeros(1, 6);
fv_sq = zeros(1, 6);
f_uv = 0;

%for u
f_u = [7 15 25 23 20 10]; % given
fu = zeros(1, 6);
fu_sq = zeros(1, 6);

% calculating fu, fv, fu^2, fv^2
for a = 1:6
    fv(a) = f_v(a) .* v_vals(a);
    fv_sq(a) = f_v(a) .* (v_vals(a)^2);
    fu(a) = f_u(a) .* u_vals(a);
    fu_sq(a) = f_u(a) .* (u_vals(a)^2);
end

% calculating f_uv
for m = 1:36
    f_uv = f_uv + fuv(m);
end
% calculating respective sums
fv_sum = 0;
fv_sq_sum = 0;
fu_sum = 0;
fu_sq_sum = 0;

for d = 1:6
    fv_sum =fv_sum + fv(d);
    fv_sq_sum =fv_sq_sum + fv_sq(d);
    fu_sum =fu_sum + fu(d);
    fu_sq_sum =fu_sq_sum + fu_sq(d);
end

% calculating square of summation of fu & fv respectively
fu_sum_sq = fu_sum ^ 2;
fv_sum_sq = fv_sum ^ 2;

% calculating correlation coefficient between x & y
corr_coeff_num = (n*f_uv) - (fu_sum*fv_sum);
corr_coeff_denom = (sqrt((n*fu_sq_sum)-fu_sum_sq))*(sqrt((n*fv_sq_sum)-
fv_sum_sq));
corr_coeff = corr_coeff_num / corr_coeff_denom;

fprintf('Correlation coefficient between the marks in Mathematics and
the marks in Physics is: %.4f\n', corr_coeff);
```

Q9.

```
clc, clearvars, close all

X = [56 42 72 36 63 47 55 49 38 42 68 60];
Y = [147 125 160 118 149 128 150 145 115 140 152 155];

n = length(Y);
Y_sq = Y .^ 2;
X_sq = X .^ 2;
YX = zeros(1, n);

for i = 1:n
    YX(i) = Y(i) * X(i);
end

Y_sum = 0;
X_sum = 0;
Y_sq_sum = 0;
X_sq_sum = 0;
YX_sum = 0;

for j = 1:n
    Y_sum = Y_sum + Y(j);
    X_sum = X_sum + X(j);
    Y_sq_sum = Y_sq_sum + Y_sq(j);
    X_sq_sum = X_sq_sum + X_sq(j);
    YX_sum = YX_sum + YX(j);
end

a_num = (n * YX_sum) - (Y_sum * X_sum);
a_denom = (n * Y_sq_sum) - (Y_sum ^ 2);

b_num = (X_sum * Y_sq_sum) - (Y_sum * YX_sum);
b_denom = (n * Y_sq_sum) - (Y_sum ^ 2);

a_eqn = a_num / a_denom;
b_eqn = b_num / b_denom;

fprintf('The regression equation is X = %.2fY + %.2f\n', a_eqn, b_eqn);

%
% Plotting the data points
scatter(Y, X, 'filled');
hold on;

% Plotting the regression line
y_line = min(Y):0.01:max(Y);
```

```
x_line = a_eqn * y_line + b_eqn;
plot(y_line, x_line, 'r');

% Formatting the plot
xlabel('Y');
ylabel('X');
title('Regression of X on Y');
legend('Data points', 'Regression line');
grid on;
hold off;
```

Q9_ALT.

```
clc, clearvars, close all

x = [56 42 72 36 63 47 55 49 38 42 68 60];
y = [147 125 160 118 149 128 150 145 115 140 152 155];
n = length(x);

sumx = 0;
sumy = 0;

for i = 1:n
    sumx = sumx + x(i);
    sumy = sumy + y(i);
end

sumxx = sum(x .^ 2);
sumyy = sum(y .^ 2);
sumxy = sum(x .* y);
mean_x = sumx / n;
mean_y = sumy / n;
Sx = n * (sumxy) - ((sumx) * (sumy));
Sy = n * (sumyy) - (sumy) ^ 2;
bxy = Sx / Sy;
x = mean_x + bxy * (y - mean_y);

fprintf('Equation of the given regression line of x on y is: \n');
fprintf('x-%f=%f(y-%f) \n',mean_x,bxy,mean_y);
plot(x, y, 'm-*')
title('REGRESSION LINE OF X ON Y')
```

Q10.

```
clc, clearvars, close all

X = [56 42 72 36 63 47 55 49 38 42 68 60];
Y = [147 125 160 118 149 128 150 145 115 140 152 155];

n = length(X);
X_sq = X .^ 2;
Y_sq = Y .^ 2;
XY = zeros(1, n);

for i = 1:n
    XY(i) = X(i) * Y(i);
end

X_sum = 0;
Y_sum = 0;
X_sq_sum = 0;
Y_sq_sum = 0;
XY_sum = 0;

for j = 1:n
    X_sum = X_sum + X(j);
    Y_sum = Y_sum + Y(j);
    X_sq_sum = X_sq_sum + X_sq(j);
    Y_sq_sum = Y_sq_sum + Y_sq(j);
    XY_sum = XY_sum + XY(j);
end

a_num = (n * XY_sum) - (X_sum * Y_sum);
a_denom = (n * X_sq_sum) - (X_sum ^ 2);

b_num = (Y_sum * X_sq_sum) - (X_sum * XY_sum);
b_denom = (n * X_sq_sum) - (X_sum ^ 2);

a_eqn = a_num / a_denom;
b_eqn = b_num / b_denom;

fprintf('The regression equation is Y = %.2fX + %.2f\n', a_eqn, b_eqn);

%
% Plotting the data points
scatter(X, Y, 'filled');
hold on;

% Plotting the regression line
x_line = min(X):0.01:max(X);
```

```
y_line = a_eqn * x_line + b_eqn;
plot(x_line, y_line, 'r');

% Formatting the plot
xlabel('X');
ylabel('Y');
title('Regression of Y on X');
legend('Data points', 'Regression line');
grid on;
hold off;
```

Q10_ALT.

```
clc, clearvars, close all

x = [56 42 72 36 63 47 55 49 38 42 68 60];
y = [147 125 160 118 149 128 150 145 115 140 152 155];
n = length(x);

sumx = 0;
sumy = 0;

for i = 1:n
    sumx = sumx+x(i);
    sumy = sumy+y(i);
end

sumxx = sum(x .^ 2);
sumyy = sum(y .^ 2);
sumxy = sum(x .* y);
mean_x = sumx / n;
mean_y = sumy / n;
Sx = n * (sumxy) - ((sumx) * (sumy));
Sy = n * (sumxx) - (sumx) ^ 2;
byx = Sx / Sy;
y = mean_y + byx * (x-mean_x);

fprintf('Equation of the given regression line of y on x is: \n');
fprintf('y-%f=%f(x-%f) \n',mean_y,byx,mean_x);
plot(x,y,'m-*')
title('REGRESSION LINE OF Y ON X')
```

Q11.

```
clc, clearvars, close all

math_x = [43 77 64 96 48 35 86 71];
phys_y = [36 68 49 79 50 41 82 65];
math_x_cp = math_x;
phys_y_cp = phys_y;
n = length(math_x);

R_x = zeros(1, n);
R_y = zeros(1, n);

for i = 1:n
    [valx, idx] = max(math_x);
    math_x(idx) = -Inf;
    R_x(idx) = i;

    [valy, idx] = max(phys_y);
    phys_y(idx) = -Inf;
    R_y(idx) = i;
end

d_sq = (R_x - R_y) .^ 2;
d_sq_sum = sum(d_sq);

sp_num = 6 * d_sq_sum;
sp_denom = n * ((n ^ 2) - 1);

sp = 1 - (sp_num / sp_denom);
table_t = zeros(8,5);
table_t(1:8, 1) = math_x_cp;
table_t(1:8, 2) = phys_y_cp;
table_t(1:8, 3) = R_x;
table_t(1:8, 4) = R_y;
table_t(1:8, 5) = d_sq;
disp(table_t)
fprintf('The Spearman Rank Correlation is %.4f\n', sp);
```

Q12.

```
clc, clearvars, close all

math_x = [43 77 64 96 48 35 86 71];
math_x_cp = math_x;
stat_y = [41 68 50 82 49 36 79 65];
stat_y_cp = stat_y;
n = length(math_x);

R_x = zeros(1, n);
R_y = zeros(1, n);

for i = 1:n
    [valx, idx] = max(math_x);
    math_x(idx) = -Inf;
    R_x(idx) = i;

    [valy, idx] = max(stat_y);
    stat_y(idx) = -Inf;
    R_y(idx) = i;
end

d_sq = (R_x - R_y) .^ 2;
d_sq_sum = sum(d_sq);

sp_num = 6 * d_sq_sum;
sp_denom = n * ((n ^ 2) - 1);

sp = 1 - (sp_num / sp_denom);

table_t = zeros(8,5);
table_t(1:8, 1) = math_x_cp;
table_t(1:8, 2) = stat_y_cp;
table_t(1:8, 3) = R_x;
table_t(1:8, 4) = R_y;
table_t(1:8, 5) = d_sq;
disp(table_t)
fprintf('The Spearman Rank Correlation is %.4f\n', sp);
```

Q13.

```
clc, clearvars, close all

math_x = [43 77 64 96 48 35 86 71];
beng_y = [79 49 65 36 68 82 41 50];
math_x_cp = math_x;
beng_y_cp = beng_y;
n = length(math_x);

R_x = zeros(1, n);
R_y = zeros(1, n);

for i = 1:n
    [valx, idx] = max(math_x);
    math_x(idx) = -Inf;
    R_x(idx) = i;

    [valy, idx] = max(beng_y);
    beng_y(idx) = -Inf;
    R_y(idx) = i;
end

d_sq = (R_x - R_y) .^ 2;
d_sq_sum = sum(d_sq);

sp_num = 6 * d_sq_sum;
sp_denom = n * ((n ^ 2) - 1);

sp = 1 - (sp_num / sp_denom);
table_t = zeros(8,5);
table_t(1:8, 1) = math_x_cp;
table_t(1:8, 2) = beng_y_cp;
table_t(1:8, 3) = R_x;
table_t(1:8, 4) = R_y;
table_t(1:8, 5) = d_sq;
disp(table_t)
fprintf('The Spearman Rank Correlation is %.4f\n', sp);
```