DMT Lab 4:

Airquality

Summary(airquality)

Mean(airquality$Ozone)

mean(airquality$Ozone)

[1] NA

> mean(airquality$ozone,na.rm=True)

[1] NA

|  |
| --- |
| mean(airquality$Ozone,na.rm=TRUE)  [1] 42.12931  > max(airquality$Ozone,na.rm=TRUE)  [1] 168  > max(airquality$Ozone)  [1] NA |
|  |
| |  | | --- | | >  Inorder to create duplicate data |   air=airquality  > air  air  Ozone Solar.R Wind Temp Month Day  1 41 190 7.4 67 5 1  2 36 118 8.0 72 5 2  3 12 149 12.6 74 5 3  4 18 313 11.5 62 5 4  5 NA NA 14.3 56 5 5  6 28 NA 14.9 66 5 6  7 23 299 8.6 65 5 7  8 19 99 13.8 59 5 8  9 8 19 20.1 61 5 9  10 NA 194 8.6 69 5 10  11 7 NA 6.9 74 5 11  12 16 256 9.7 69 5 12  13 11 290 9.2 66 5 13  14 14 274 10.9 68 5 14  15 18 65 13.2 58 5 15  16 14 334 11.5 64 5 16  17 34 307 12.0 66 5 17  18 6 78 18.4 57 5 18  19 30 322 11.5 68 5 19  20 11 44 9.7 62 5 20  21 1 8 9.7 59 5 21  22 11 320 16.6 73 5 22  23 4 25 9.7 61 5 23  24 32 92 12.0 61 5 24  25 NA 66 16.6 57 5 25  26 NA 266 14.9 58 5 26  27 NA NA 8.0 57 5 27  28 23 13 12.0 67 5 28  29 45 252 14.9 81 5 29  30 115 223 5.7 79 5 30  31 37 279 7.4 76 5 31  32 NA 286 8.6 78 6 1  33 NA 287 9.7 74 6 2  34 NA 242 16.1 67 6 3  35 NA 186 9.2 84 6 4  36 NA 220 8.6 85 6 5  37 NA 264 14.3 79 6 6  38 29 127 9.7 82 6 7  39 NA 273 6.9 87 6 8  40 71 291 13.8 90 6 9  41 39 323 11.5 87 6 10  42 NA 259 10.9 93 6 11  43 NA 250 9.2 92 6 12  44 23 148 8.0 82 6 13  45 NA 332 13.8 80 6 14  46 NA 322 11.5 79 6 15  47 21 191 14.9 77 6 16  48 37 284 20.7 72 6 17  49 20 37 9.2 65 6 18  50 12 120 11.5 73 6 19  51 13 137 10.3 76 6 20  52 NA 150 6.3 77 6 21  53 NA 59 1.7 76 6 22  54 NA 91 4.6 76 6 23  55 NA 250 6.3 76 6 24  56 NA 135 8.0 75 6 25  57 NA 127 8.0 78 6 26  58 NA 47 10.3 73 6 27  59 NA 98 11.5 80 6 28  60 NA 31 14.9 77 6 29  61 NA 138 8.0 83 6 30  62 135 269 4.1 84 7 1  63 49 248 9.2 85 7 2  64 32 236 9.2 81 7 3  65 NA 101 10.9 84 7 4  66 64 175 4.6 83 7 5  67 40 314 10.9 83 7 6  68 77 276 5.1 88 7 7  69 97 267 6.3 92 7 8  70 97 272 5.7 92 7 9  71 85 175 7.4 89 7 10  72 NA 139 8.6 82 7 11  73 10 264 14.3 73 7 12  74 27 175 14.9 81 7 13  75 NA 291 14.9 91 7 14  76 7 48 14.3 80 7 15  77 48 260 6.9 81 7 16  78 35 274 10.3 82 7 17  79 61 285 6.3 84 7 18  80 79 187 5.1 87 7 19  81 63 220 11.5 85 7 20  82 16 7 6.9 74 7 21  83 NA 258 9.7 81 7 22  84 NA 295 11.5 82 7 23  85 80 294 8.6 86 7 24  86 108 223 8.0 85 7 25  87 20 81 8.6 82 7 26  88 52 82 12.0 86 7 27  89 82 213 7.4 88 7 28  90 50 275 7.4 86 7 29  91 64 253 7.4 83 7 30  92 59 254 9.2 81 7 31  93 39 83 6.9 81 8 1  94 9 24 13.8 81 8 2  95 16 77 7.4 82 8 3  96 78 NA 6.9 86 8 4  97 35 NA 7.4 85 8 5  98 66 NA 4.6 87 8 6  99 122 255 4.0 89 8 7  100 89 229 10.3 90 8 8  101 110 207 8.0 90 8 9  102 NA 222 8.6 92 8 10  103 NA 137 11.5 86 8 11  104 44 192 11.5 86 8 12  105 28 273 11.5 82 8 13  106 65 157 9.7 80 8 14  107 NA 64 11.5 79 8 15  108 22 71 10.3 77 8 16  109 59 51 6.3 79 8 17  110 23 115 7.4 76 8 18  111 31 244 10.9 78 8 19  112 44 190 10.3 78 8 20  113 21 259 15.5 77 8 21  114 9 36 14.3 72 8 22  115 NA 255 12.6 75 8 23  116 45 212 9.7 79 8 24  117 168 238 3.4 81 8 25  118 73 215 8.0 86 8 26  119 NA 153 5.7 88 8 27  120 76 203 9.7 97 8 28  121 118 225 2.3 94 8 29  122 84 237 6.3 96 8 30  123 85 188 6.3 94 8 31  124 96 167 6.9 91 9 1  125 78 197 5.1 92 9 2  126 73 183 2.8 93 9 3  127 91 189 4.6 93 9 4  128 47 95 7.4 87 9 5  129 32 92 15.5 84 9 6  130 20 252 10.9 80 9 7  131 23 220 10.3 78 9 8  132 21 230 10.9 75 9 9  133 24 259 9.7 73 9 10  134 44 236 14.9 81 9 11  135 21 259 15.5 76 9 12  136 28 238 6.3 77 9 13  137 9 24 10.9 71 9 14  138 13 112 11.5 71 9 15  139 46 237 6.9 78 9 16  140 18 224 13.8 67 9 17  141 13 27 10.3 76 9 18  142 24 238 10.3 68 9 19  143 16 201 8.0 82 9 20  144 13 238 12.6 64 9 21  145 23 14 9.2 71 9 22  146 36 139 10.3 81 9 23  147 7 49 10.3 69 9 24  148 14 20 16.6 63 9 25  149 30 193 6.9 70 9 26  150 NA 145 13.2 77 9 27  151 14 191 14.3 75 9 28  152 18 131 8.0 76 9 29  153 20 223 11.5 68 9 30   |  | | --- | | head(air)  Ozone Solar.R Wind Temp Month Day  1 41 190 7.4 67 5 1  2 36 118 8.0 72 5 2  3 12 149 12.6 74 5 3  4 18 313 11.5 62 5 4  5 NA NA 14.3 56 5 5  6 28 NA 14.9 66 5 6  > tail(air)  Ozone Solar.R Wind Temp Month Day  148 14 20 16.6 63 9 25  149 30 193 6.9 70 9 26  150 NA 145 13.2 77 9 27  151 14 191 14.3 75 9 28  152 18 131 8.0 76 9 29  153 20 223 11.5 68 9 30 | |  | | |  |  |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | --- | --- | | >  Replacing missing values with median  air$Ozone=ifelse(is.na(air$Ozone),median(air$Ozone,na.rm=TRUE),air$Ozone)  > summary(air)  Ozone Solar.R Wind  Min. : 1.00 Min. : 7.0 Min. : 1.700  1st Qu.: 21.00 1st Qu.:115.8 1st Qu.: 7.400  Median : 31.50 Median :205.0 Median : 9.700  Mean : 39.56 Mean :185.9 Mean : 9.958  3rd Qu.: 46.00 3rd Qu.:258.8 3rd Qu.:11.500  Max. :168.00 Max. :334.0 Max. :20.700  NA's :7  Temp Month Day  Min. :56.00 Min. :5.000 Min. : 1.0  1st Qu.:72.00 1st Qu.:6.000 1st Qu.: 8.0  Median :79.00 Median :7.000 Median :16.0  Mean :77.88 Mean :6.993 Mean :15.8  3rd Qu.:85.00 3rd Qu.:8.000 3rd Qu.:23.0  Max. :97.00 Max. :9.000 Max. :31.0  Replacing missing values with mean  air$Ozone=ifelse(is.na(air$Ozone),mean(air$Ozone,na.rm=TRUE),air$Ozone)  > summary(air)  Ozone Solar.R Wind  Min. : 1.00 Min. : 7.0 Min. : 1.700  1st Qu.: 21.00 1st Qu.:115.8 1st Qu.: 7.400  Median : 31.50 Median :205.0 Median : 9.700  Mean : 39.56 Mean :185.9 Mean : 9.958  3rd Qu.: 46.00 3rd Qu.:258.8 3rd Qu.:11.500  Max. :168.00 Max. :334.0 Max. :20.700  NA's :7  Temp Month Day  Min. :56.00 Min. :5.000 Min. : 1.0  1st Qu.:72.00 1st Qu.:6.000 1st Qu.: 8.0  Median :79.00 Median :7.000 Median :16.0  Mean :77.88 Mean :6.993 Mean :15.8  3rd Qu.:85.00 3rd Qu.:8.000 3rd Qu.:23.0  Max. :97.00 Max. :9.000 Max. :31.0  Repklacing missing value with median in Solar.R:  air$Solar.R=ifelse(is.na(air$Solar.R),median(air$Solar.R,na.rm=TRUE),air$Solar.R)  > summary(air)  Ozone Solar.R Wind  Min. : 7.0 Min. : 7.0 Min. : 1.700  1st Qu.:120.0 1st Qu.:120.0 1st Qu.: 7.400  Median :205.0 Median :205.0 Median : 9.700  Mean :186.8 Mean :186.8 Mean : 9.958  3rd Qu.:256.0 3rd Qu.:256.0 3rd Qu.:11.500  Max. :334.0 Max. :334.0 Max. :20.700  Temp Month Day  Min. :56.00 Min. :5.000 Min. : 1.0  1st Qu.:72.00 1st Qu.:6.000 1st Qu.: 8.0  Median :79.00 Median :7.000 Median :16.0  Mean :77.88 Mean :6.993 Mean :15.8  3rd Qu.:85.00 3rd Qu.:8.000 3rd Qu.:23.0  Max. :97.00 Max. :9.000 Max. :31.0  > head(air)  Ozone Solar.R Wind Temp Month Day  1 190 190 7.4 67 5 1  2 118 118 8.0 72 5 2  3 149 149 12.6 74 5 3  4 313 313 11.5 62 5 4  5 205 205 14.3 56 5 5  6 205 205 14.9 66 5 6  > tail(air)  Ozone Solar.R Wind Temp Month Day  148 20 20 16.6 63 9 25  149 193 193 6.9 70 9 26  150 145 145 13.2 77 9 27  151 191 191 14.3 75 9 28  152 131 131 8.0 76 9 29  153 223 223 11.5 68 9 30  Replacing missing value with mean in Solar.R:  air$Solar.R=ifelse(is.na(air$Solar.R),mean(air$Solar.R,na.rm=TRUE),air$Solar.R)  > summary(air)  Ozone Solar.R Wind  Min. : 7.0 Min. : 7.0 Min. : 1.700  1st Qu.:120.0 1st Qu.:120.0 1st Qu.: 7.400  Median :205.0 Median :205.0 Median : 9.700  Mean :186.8 Mean :186.8 Mean : 9.958  3rd Qu.:256.0 3rd Qu.:256.0 3rd Qu.:11.500  Max. :334.0 Max. :334.0 Max. :20.700  Temp Month Day  Min. :56.00 Min. :5.000 Min. : 1.0  1st Qu.:72.00 1st Qu.:6.000 1st Qu.: 8.0  Median :79.00 Median :7.000 Median :16.0  Mean :77.88 Mean :6.993 Mean :15.8  3rd Qu.:85.00 3rd Qu.:8.000 3rd Qu.:23.0  Max. :97.00 Max. :9.000 Max. :31.0  head(air)  Ozone Solar.R Wind Temp Month Day  1 190 190 7.4 67 5 1  2 118 118 8.0 72 5 2  3 149 149 12.6 74 5 3  4 313 313 11.5 62 5 4  5 205 205 14.3 56 5 5  6 205 205 14.9 66 5 6  > tail(air)  Ozone Solar.R Wind Temp Month Day  148 20 20 16.6 63 9 25  149 193 193 6.9 70 9 26  150 145 145 13.2 77 9 27  151 191 191 14.3 75 9 28  152 131 131 8.0 76 9 29  153 223 223 11.5 68 9 30  Normalization based on binning:  brks=c(0,50,100,150,200,250,300,350)  # dividing based on four groups  > brks  [1] 0 50 100 150 200 250 300 350  > air$Solar.R=cut(air$Solar.R,breaks=brks,include.lowest = TRUE)  > head(air)  Ozone Solar.R Wind Temp Month Day  1 190 (150,200] 7.4 67 5 1  2 118 (100,150] 8.0 72 5 2  3 149 (100,150] 12.6 74 5 3  4 313 (300,350] 11.5 62 5 4  5 205 (200,250] 14.3 56 5 5  6 205 (200,250] 14.9 66 5 6  Replacing month number with month name:  >Air1=airquality  >air1$Month=gsub(5,"May",air1$Month)  > air1$Month=gsub(6,"June",air1$Month)  > air1$Month=gsub(7,"July",air1$Month)  > air1$Month=gsub(8,"Aug",air1$Month)  > air1$Month=gsub(9,"Sep",air1$Month)  > air1   |  | | --- | | Ozone Solar.R Wind Temp Month Day  1 41 190 7.4 67 May 1  2 36 118 8.0 72 May 2  3 12 149 12.6 74 May 3  4 18 313 11.5 62 May 4  5 NA NA 14.3 56 May 5  6 28 NA 14.9 66 May 6  7 23 299 8.6 65 May 7  8 19 99 13.8 59 May 8  9 8 19 20.1 61 May 9  10 NA 194 8.6 69 May 10  11 7 NA 6.9 74 May 11  12 16 256 9.7 69 May 12  13 11 290 9.2 66 May 13  14 14 274 10.9 68 May 14  15 18 65 13.2 58 May 15  16 14 334 11.5 64 May 16  17 34 307 12.0 66 May 17  18 6 78 18.4 57 May 18  19 30 322 11.5 68 May 19  20 11 44 9.7 62 May 20  21 1 8 9.7 59 May 21  22 11 320 16.6 73 May 22  23 4 25 9.7 61 May 23  24 32 92 12.0 61 May 24  25 NA 66 16.6 57 May 25  26 NA 266 14.9 58 May 26  27 NA NA 8.0 57 May 27  28 23 13 12.0 67 May 28  29 45 252 14.9 81 May 29  30 115 223 5.7 79 May 30  31 37 279 7.4 76 May 31  32 NA 286 8.6 78 June 1  33 NA 287 9.7 74 June 2  34 NA 242 16.1 67 June 3  35 NA 186 9.2 84 June 4  36 NA 220 8.6 85 June 5  37 NA 264 14.3 79 June 6  38 29 127 9.7 82 June 7  39 NA 273 6.9 87 June 8  40 71 291 13.8 90 June 9  41 39 323 11.5 87 June 10  42 NA 259 10.9 93 June 11  43 NA 250 9.2 92 June 12  44 23 148 8.0 82 June 13  45 NA 332 13.8 80 June 14  46 NA 322 11.5 79 June 15  47 21 191 14.9 77 June 16  48 37 284 20.7 72 June 17  49 20 37 9.2 65 June 18  50 12 120 11.5 73 June 19  51 13 137 10.3 76 June 20  52 NA 150 6.3 77 June 21  53 NA 59 1.7 76 June 22  54 NA 91 4.6 76 June 23  55 NA 250 6.3 76 June 24  56 NA 135 8.0 75 June 25  57 NA 127 8.0 78 June 26  58 NA 47 10.3 73 June 27  59 NA 98 11.5 80 June 28  60 NA 31 14.9 77 June 29  61 NA 138 8.0 83 June 30  62 135 269 4.1 84 July 1  63 49 248 9.2 85 July 2  64 32 236 9.2 81 July 3  65 NA 101 10.9 84 July 4  66 64 175 4.6 83 July 5  67 40 314 10.9 83 July 6  68 77 276 5.1 88 July 7  69 97 267 6.3 92 July 8  70 97 272 5.7 92 July 9  71 85 175 7.4 89 July 10  72 NA 139 8.6 82 July 11  73 10 264 14.3 73 July 12  74 27 175 14.9 81 July 13  75 NA 291 14.9 91 July 14  76 7 48 14.3 80 July 15  77 48 260 6.9 81 July 16  78 35 274 10.3 82 July 17  79 61 285 6.3 84 July 18  80 79 187 5.1 87 July 19  81 63 220 11.5 85 July 20  82 16 7 6.9 74 July 21  83 NA 258 9.7 81 July 22  84 NA 295 11.5 82 July 23  85 80 294 8.6 86 July 24  86 108 223 8.0 85 July 25  87 20 81 8.6 82 July 26  88 52 82 12.0 86 July 27  89 82 213 7.4 88 July 28  90 50 275 7.4 86 July 29  91 64 253 7.4 83 July 30  92 59 254 9.2 81 July 31  93 39 83 6.9 81 Aug 1  94 9 24 13.8 81 Aug 2  95 16 77 7.4 82 Aug 3  96 78 NA 6.9 86 Aug 4  97 35 NA 7.4 85 Aug 5  98 66 NA 4.6 87 Aug 6  99 122 255 4.0 89 Aug 7  100 89 229 10.3 90 Aug 8  101 110 207 8.0 90 Aug 9  102 NA 222 8.6 92 Aug 10  103 NA 137 11.5 86 Aug 11  104 44 192 11.5 86 Aug 12  105 28 273 11.5 82 Aug 13  106 65 157 9.7 80 Aug 14  107 NA 64 11.5 79 Aug 15  108 22 71 10.3 77 Aug 16  109 59 51 6.3 79 Aug 17  110 23 115 7.4 76 Aug 18  111 31 244 10.9 78 Aug 19  112 44 190 10.3 78 Aug 20  113 21 259 15.5 77 Aug 21  114 9 36 14.3 72 Aug 22  115 NA 255 12.6 75 Aug 23  116 45 212 9.7 79 Aug 24  117 168 238 3.4 81 Aug 25  118 73 215 8.0 86 Aug 26  119 NA 153 5.7 88 Aug 27  120 76 203 9.7 97 Aug 28  121 118 225 2.3 94 Aug 29  122 84 237 6.3 96 Aug 30  123 85 188 6.3 94 Aug 31  124 96 167 6.9 91 Sep 1  125 78 197 5.1 92 Sep 2  126 73 183 2.8 93 Sep 3  127 91 189 4.6 93 Sep 4  128 47 95 7.4 87 Sep 5  129 32 92 15.5 84 Sep 6  130 20 252 10.9 80 Sep 7  131 23 220 10.3 78 Sep 8  132 21 230 10.9 75 Sep 9  133 24 259 9.7 73 Sep 10  134 44 236 14.9 81 Sep 11  135 21 259 15.5 76 Sep 12  136 28 238 6.3 77 Sep 13  137 9 24 10.9 71 Sep 14  138 13 112 11.5 71 Sep 15  139 46 237 6.9 78 Sep 16  140 18 224 13.8 67 Sep 17  141 13 27 10.3 76 Sep 18  142 24 238 10.3 68 Sep 19  143 16 201 8.0 82 Sep 20  144 13 238 12.6 64 Sep 21  145 23 14 9.2 71 Sep 22  146 36 139 10.3 81 Sep 23  147 7 49 10.3 69 Sep 24  148 14 20 16.6 63 Sep 25  149 30 193 6.9 70 Sep 26  150 NA 145 13.2 77 Sep 27  151 14 191 14.3 75 Sep 28  152 18 131 8.0 76 Sep 29  153 20 223 11.5 68 Sep 30 | |  | | |  |  |  |  |  | | --- | --- | --- | --- | --- | | >  Normalizing data==after data cleaning:  In this case I’m replacing with mean values   |  | | --- | | air1$Ozone=ifelse(is.na(air1$Ozone),mean(air1$Ozone,na.rm=TRUE),air1$Ozone)  > normalize<- function(x){  + return((x-min(x))/(max(x)-min(x)))  + }  > normalize(air1$Ozone)  [1] 0.23952096 0.20958084 0.06586826 0.10179641 0.24628330  [6] 0.16167665 0.13173653 0.10778443 0.04191617 0.24628330  [11] 0.03592814 0.08982036 0.05988024 0.07784431 0.10179641  [16] 0.07784431 0.19760479 0.02994012 0.17365269 0.05988024  [21] 0.00000000 0.05988024 0.01796407 0.18562874 0.24628330  [26] 0.24628330 0.24628330 0.13173653 0.26347305 0.68263473  [31] 0.21556886 0.24628330 0.24628330 0.24628330 0.24628330  [36] 0.24628330 0.24628330 0.16766467 0.24628330 0.41916168  [41] 0.22754491 0.24628330 0.24628330 0.13173653 0.24628330  [46] 0.24628330 0.11976048 0.21556886 0.11377246 0.06586826  [51] 0.07185629 0.24628330 0.24628330 0.24628330 0.24628330  [56] 0.24628330 0.24628330 0.24628330 0.24628330 0.24628330  [61] 0.24628330 0.80239521 0.28742515 0.18562874 0.24628330  [66] 0.37724551 0.23353293 0.45508982 0.57485030 0.57485030  [71] 0.50299401 0.24628330 0.05389222 0.15568862 0.24628330  [76] 0.03592814 0.28143713 0.20359281 0.35928144 0.46706587  [81] 0.37125749 0.08982036 0.24628330 0.24628330 0.47305389  [86] 0.64071856 0.11377246 0.30538922 0.48502994 0.29341317  [91] 0.37724551 0.34730539 0.22754491 0.04790419 0.08982036  [96] 0.46107784 0.20359281 0.38922156 0.72455090 0.52694611  [101] 0.65269461 0.24628330 0.24628330 0.25748503 0.16167665  [106] 0.38323353 0.24628330 0.12574850 0.34730539 0.13173653  [111] 0.17964072 0.25748503 0.11976048 0.04790419 0.24628330  [116] 0.26347305 1.00000000 0.43113772 0.24628330 0.44910180  [121] 0.70059880 0.49700599 0.50299401 0.56886228 0.46107784  [126] 0.43113772 0.53892216 0.27544910 0.18562874 0.11377246  [131] 0.13173653 0.11976048 0.13772455 0.25748503 0.11976048  [136] 0.16167665 0.04790419 0.07185629 0.26946108 0.10179641  [141] 0.07185629 0.13772455 0.08982036 0.07185629 0.13173653  [146] 0.20958084 0.03592814 0.07784431 0.17365269 0.24628330  [151] 0.07784431 0.10179641 0.11377246 | |  | | |  | | --- | | > | | | | | |   **Normalization based on month:**  If condition====we take particular attribute instead of entire attribute value I.e I will consider the mean of Ozone attributr  **Ozone attribute ne month attribute tho normalize**  air3=airquality  > for (i in 1:nrow(air3)){  + if(is.na(air3[i,"Ozone"])){  + air3[i,"Ozone"]<- mean(air3[which(air3[,"Month"]==air3[i,"Month"]),"Ozone"],na.rm = TRUE)  + }  +  + }  > summary(air3)  Ozone Solar.R Wind Temp Month Day  1 41.00000 190 7.4 67 5 1  2 36.00000 118 8.0 72 5 2  3 12.00000 149 12.6 74 5 3  4 18.00000 313 11.5 62 5 4  5 23.61538 NA 14.3 56 5 5  6 28.00000 NA 14.9 66 5 6  7 23.00000 299 8.6 65 5 7  8 19.00000 99 13.8 59 5 8  9 8.00000 19 20.1 61 5 9  10 23.61538 194 8.6 69 5 10  11 7.00000 NA 6.9 74 5 11  12 16.00000 256 9.7 69 5 12  13 11.00000 290 9.2 66 5 13  14 14.00000 274 10.9 68 5 14  15 18.00000 65 13.2 58 5 15  16 14.00000 334 11.5 64 5 16  17 34.00000 307 12.0 66 5 17  18 6.00000 78 18.4 57 5 18  19 30.00000 322 11.5 68 5 19  20 11.00000 44 9.7 62 5 20  21 1.00000 8 9.7 59 5 21  22 11.00000 320 16.6 73 5 22  23 4.00000 25 9.7 61 5 23  24 32.00000 92 12.0 61 5 24  25 23.61538 66 16.6 57 5 25  26 23.61538 266 14.9 58 5 26  27 23.61538 NA 8.0 57 5 27  28 23.00000 13 12.0 67 5 28  29 45.00000 252 14.9 81 5 29  30 115.00000 223 5.7 79 5 30  31 37.00000 279 7.4 76 5 31  32 29.44444 286 8.6 78 6 1  33 29.44444 287 9.7 74 6 2  34 29.44444 242 16.1 67 6 3  35 29.44444 186 9.2 84 6 4  36 29.44444 220 8.6 85 6 5  37 29.44444 264 14.3 79 6 6  38 29.00000 127 9.7 82 6 7  39 29.44444 273 6.9 87 6 8  40 71.00000 291 13.8 90 6 9  41 39.00000 323 11.5 87 6 10  42 29.44444 259 10.9 93 6 11  43 29.44444 250 9.2 92 6 12  44 23.00000 148 8.0 82 6 13  45 29.44444 332 13.8 80 6 14  46 29.44444 322 11.5 79 6 15  47 21.00000 191 14.9 77 6 16  48 37.00000 284 20.7 72 6 17  49 20.00000 37 9.2 65 6 18  50 12.00000 120 11.5 73 6 19  51 13.00000 137 10.3 76 6 20  52 29.44444 150 6.3 77 6 21  53 29.44444 59 1.7 76 6 22  54 29.44444 91 4.6 76 6 23  55 29.44444 250 6.3 76 6 24  56 29.44444 135 8.0 75 6 25  57 29.44444 127 8.0 78 6 26  58 29.44444 47 10.3 73 6 27  59 29.44444 98 11.5 80 6 28  60 29.44444 31 14.9 77 6 29  61 29.44444 138 8.0 83 6 30  62 135.00000 269 4.1 84 7 1  63 49.00000 248 9.2 85 7 2  64 32.00000 236 9.2 81 7 3  65 59.11538 101 10.9 84 7 4  66 64.00000 175 4.6 83 7 5  67 40.00000 314 10.9 83 7 6  68 77.00000 276 5.1 88 7 7  69 97.00000 267 6.3 92 7 8  70 97.00000 272 5.7 92 7 9  71 85.00000 175 7.4 89 7 10  72 59.11538 139 8.6 82 7 11  73 10.00000 264 14.3 73 7 12  74 27.00000 175 14.9 81 7 13  75 59.11538 291 14.9 91 7 14  76 7.00000 48 14.3 80 7 15  77 48.00000 260 6.9 81 7 16  78 35.00000 274 10.3 82 7 17  79 61.00000 285 6.3 84 7 18  80 79.00000 187 5.1 87 7 19  81 63.00000 220 11.5 85 7 20  82 16.00000 7 6.9 74 7 21  83 59.11538 258 9.7 81 7 22  84 59.11538 295 11.5 82 7 23  85 80.00000 294 8.6 86 7 24  86 108.00000 223 8.0 85 7 25  87 20.00000 81 8.6 82 7 26  88 52.00000 82 12.0 86 7 27  89 82.00000 213 7.4 88 7 28  90 50.00000 275 7.4 86 7 29  91 64.00000 253 7.4 83 7 30  92 59.00000 254 9.2 81 7 31  93 39.00000 83 6.9 81 8 1  94 9.00000 24 13.8 81 8 2  95 16.00000 77 7.4 82 8 3  96 78.00000 NA 6.9 86 8 4  97 35.00000 NA 7.4 85 8 5  98 66.00000 NA 4.6 87 8 6  99 122.00000 255 4.0 89 8 7  100 89.00000 229 10.3 90 8 8  101 110.00000 207 8.0 90 8 9  102 59.96154 222 8.6 92 8 10  103 59.96154 137 11.5 86 8 11  104 44.00000 192 11.5 86 8 12  105 28.00000 273 11.5 82 8 13  106 65.00000 157 9.7 80 8 14  107 59.96154 64 11.5 79 8 15  108 22.00000 71 10.3 77 8 16  109 59.00000 51 6.3 79 8 17  110 23.00000 115 7.4 76 8 18  111 31.00000 244 10.9 78 8 19  112 44.00000 190 10.3 78 8 20  113 21.00000 259 15.5 77 8 21  114 9.00000 36 14.3 72 8 22  115 59.96154 255 12.6 75 8 23  116 45.00000 212 9.7 79 8 24  117 168.00000 238 3.4 81 8 25  118 73.00000 215 8.0 86 8 26  119 59.96154 153 5.7 88 8 27  120 76.00000 203 9.7 97 8 28  121 118.00000 225 2.3 94 8 29  122 84.00000 237 6.3 96 8 30  123 85.00000 188 6.3 94 8 31  124 96.00000 167 6.9 91 9 1  125 78.00000 197 5.1 92 9 2  126 73.00000 183 2.8 93 9 3  127 91.00000 189 4.6 93 9 4  128 47.00000 95 7.4 87 9 5  129 32.00000 92 15.5 84 9 6  130 20.00000 252 10.9 80 9 7  131 23.00000 220 10.3 78 9 8  132 21.00000 230 10.9 75 9 9  133 24.00000 259 9.7 73 9 10  134 44.00000 236 14.9 81 9 11  135 21.00000 259 15.5 76 9 12  136 28.00000 238 6.3 77 9 13  137 9.00000 24 10.9 71 9 14  138 13.00000 112 11.5 71 9 15  139 46.00000 237 6.9 78 9 16  140 18.00000 224 13.8 67 9 17  141 13.00000 27 10.3 76 9 18  142 24.00000 238 10.3 68 9 19  143 16.00000 201 8.0 82 9 20  144 13.00000 238 12.6 64 9 21  145 23.00000 14 9.2 71 9 22  146 36.00000 139 10.3 81 9 23  147 7.00000 49 10.3 69 9 24  148 14.00000 20 16.6 63 9 25  149 30.00000 193 6.9 70 9 26  150 31.44828 145 13.2 77 9 27  151 14.00000 191 14.3 75 9 28  152 18.00000 131 8.0 76 9 29  153 20.00000 223 11.5 68 9 30  **Regression analysis:**   |  | | --- | | Y<- air[,"Ozone"]  > X<- air[,"Solar.R"]  > model1<- lm(Y~X)  > model1  Call:  lm(formula = Y ~ X)  Coefficients:  (Intercept) X  18.5987 0.1272  > plot(Y~X)  > abline(model1, col="blue", lwd=3) | |  | | |  | | --- | | > last arguments tell about line width | |   o/p: Diagrams which shows the regression analysis.  Regression analysis on temperature and month:  Y<- air[,"Temp"]  > X<- air[,"Month"]  > model1<- lm(Y~X)  > plot(Y~X)  > abline(model1, col="blue", lwd=3)  o/p:Diagram  C:\Users\Vignan\Desktop\Rplot.png |

**Data Integration:**

sport=c("Hockey","Basketball","Football")

> leage=c("L1","L2","L3")

> tropy=c("Sachin","saurav","Yuvaraj")

trophies1=cbind(sport,leage,tropy)

> trophies1

sport leage tropy

[1,] "Hockey" "L1" "Sachin"

[2,] "Basketball" "L2" "saurav"

[3,] "Football" "L3" "Yuvaraj"

> sport1=c("Swimming","running")

> leage1=c

> leage1=c("L4","L5")

> tropy1=c("Gavaskar","Tendulkar")

trophie2=cbind(sport1,leage1,tropy1)

> trophie2

sport1 leage1 tropy1

[1,] "Swimming" "L4" "Gavaskar"

[2,] "running" "L5" "Tendulkar"