

ICA4 Module 4

word_document: default pdf_document: default —

Author: Andres Felipe Alba Hernández

Department: Electrical Engineering

Date: September 23, 2018

Course: ISYE670 Data Science for Engineers

Professor: Dr. Christine Nguyen

Northern Illinois University

```
#Initializing
rm(list=ls()) #Deleting all variables in the enviroment
#install.packages("tidyr", lib="/home/leasanspy/DataScience_NIU/Rpackages") #Installing the library
library(tidyr, lib.loc="/home/leasanspy/DataScience_NIU/Rpackages") #Calling the library
```

1. Using R functions, what are the header names?

The names are the ones listed below:

```
rev_hist=read.csv("Revenue_Historical_2000.csv", header = TRUE)
names(rev_hist)

## [1] "Year"      "Month.1"   "Month.2"   "Month.3"   "Month.4"   "Month.5"
## [7] "Month.6"   "Month.7"   "Month.8"   "Month.9"   "Month.10"  "Month.11"
## [13] "Month.12"
```

2. Output a summary of the data structure. What is the mean of the values in Month 2?

The mean of the month two is: 5558

```
#str(rev_hist)
summary(rev_hist)
```

##	Year	Month.1	Month.2	Month.3
##	Min. :2000	Min. :1008	Min. : 1119	Min. : 1204
##	1st Qu.:2004	1st Qu.:3301	1st Qu.: 3504	1st Qu.: 3652
##	Median :2008	Median :5319	Median : 5476	Median : 5643
##	Mean :2008	Mean :5405	Mean : 5558	Mean : 5716
##	3rd Qu.:2011	3rd Qu.:7516	3rd Qu.: 7628	3rd Qu.: 7844
##	Max. :2015	Max. :9950	Max. :10186	Max. :10361
##	Month.4	Month.5	Month.6	Month.7
##	Min. : 1292	Min. : 1560	Min. : 2007	Min. : 2496
##	1st Qu.: 3795	1st Qu.: 4067	1st Qu.: 4451	1st Qu.: 4731
##	Median : 5800	Median : 6030	Median : 6252	Median : 6607
##	Mean : 5866	Mean : 6190	Mean : 6520	Mean : 6866
##	3rd Qu.: 8032	3rd Qu.: 8346	3rd Qu.: 8781	3rd Qu.: 9032
##	Max. :10552	Max. :10863	Max. :11115	Max. :11545
##	Month.8	Month.9	Month.10	Month.11
##	Min. : 2696	Min. : 2255	Min. : 2022	Min. : 1817
##	1st Qu.: 5134	1st Qu.: 4724	1st Qu.: 4516	1st Qu.: 4297
##	Median : 6856	Median : 6442	Median : 6228	Median : 6012
##	Mean : 7186	Mean : 6770	Mean : 6548	Mean : 6322
##	3rd Qu.: 9334	3rd Qu.: 8842	3rd Qu.: 8622	3rd Qu.: 8386
##	Max. :11777	Max. :11405	Max. :11168	Max. :10922
##	Month.12			
##	Min. : 1594			

```
## 1st Qu.: 4070
## Median : 5774
## Mean   : 6094
## 3rd Qu.: 8163
## Max.   :10721
```

3. Use the gather function to create a new table with the desire structure:

```
#gather(data,key: out_name_column, value: out_data_name, list_of_names_with_columns_gathered)
new_table_gather=gather(rev_hist, Month, Revenue,Month.1:Month.12)
#new_table_gather
```

4. Use the separate function to separate the values under the “Month” column in part 3.

```
#names(new_table_gather)
#separate( data, col, into, sep = [regular expression], remove = TRUE, convert = FALSE)
new_table_separate=separate(new_table_gather,Month,c("Time", "Time_Interval"),remove=TRUE,convert=FALSE)
new_table_separate
```

##	Year	Time	Time_Interval	Revenue
## 1	2000	Month	1	1008
## 2	2001	Month	1	1472
## 3	2002	Month	1	2133
## 4	2003	Month	1	2844
## 5	2004	Month	1	3453
## 6	2005	Month	1	4034
## 7	2006	Month	1	4548
## 8	2007	Month	1	5039
## 9	2008	Month	1	5599
## 10	2009	Month	1	6217
## 11	2010	Month	1	6834
## 12	2011	Month	1	7357
## 13	2012	Month	1	7991
## 14	2013	Month	1	8695
## 15	2014	Month	1	9305
## 16	2015	Month	1	9950
## 17	2000	Month	2	1119
## 18	2001	Month	2	1657
## 19	2002	Month	2	2307
## 20	2003	Month	2	3011
## 21	2004	Month	2	3668
## 22	2005	Month	2	4201
## 23	2006	Month	2	4646
## 24	2007	Month	2	5165
## 25	2008	Month	2	5788
## 26	2009	Month	2	6372
## 27	2010	Month	2	6965
## 28	2011	Month	2	7464
## 29	2012	Month	2	8119
## 30	2013	Month	2	8816
## 31	2014	Month	2	9437
## 32	2015	Month	2	10186
## 33	2000	Month	3	1204
## 34	2001	Month	3	1844
## 35	2002	Month	3	2510
## 36	2003	Month	3	3155

## 37	2004 Month	3	3817
## 38	2005 Month	3	4276
## 39	2006 Month	3	4733
## 40	2007 Month	3	5367
## 41	2008 Month	3	5919
## 42	2009 Month	3	6596
## 43	2010 Month	3	7054
## 44	2011 Month	3	7695
## 45	2012 Month	3	8291
## 46	2013 Month	3	9027
## 47	2014 Month	3	9604
## 48	2015 Month	3	10361
## 49	2000 Month	4	1292
## 50	2001 Month	4	1962
## 51	2002 Month	4	2668
## 52	2003 Month	4	3286
## 53	2004 Month	4	3965
## 54	2005 Month	4	4392
## 55	2006 Month	4	4817
## 56	2007 Month	4	5494
## 57	2008 Month	4	6107
## 58	2009 Month	4	6702
## 59	2010 Month	4	7263
## 60	2011 Month	4	7874
## 61	2012 Month	4	8506
## 62	2013 Month	4	9140
## 63	2014 Month	4	9829
## 64	2015 Month	4	10552
## 65	2000 Month	5	1560
## 66	2001 Month	5	2347
## 67	2002 Month	5	2909
## 68	2003 Month	5	3739
## 69	2004 Month	5	4176
## 70	2005 Month	5	4743
## 71	2006 Month	5	5163
## 72	2007 Month	5	5701
## 73	2008 Month	5	6360
## 74	2009 Month	5	7077
## 75	2010 Month	5	7721
## 76	2011 Month	5	8155
## 77	2012 Month	5	8919
## 78	2013 Month	5	9343
## 79	2014 Month	5	10259
## 80	2015 Month	5	10863
## 81	2000 Month	6	2007
## 82	2001 Month	6	2640
## 83	2002 Month	6	3279
## 84	2003 Month	6	4024
## 85	2004 Month	6	4593
## 86	2005 Month	6	5186
## 87	2006 Month	6	5635
## 88	2007 Month	6	5925
## 89	2008 Month	6	6579
## 90	2009 Month	6	7305

## 91	2010	Month	6	7989
## 92	2011	Month	6	8634
## 93	2012	Month	6	9222
## 94	2013	Month	6	9550
## 95	2014	Month	6	10638
## 96	2015	Month	6	11115
## 97	2000	Month	7	2496
## 98	2001	Month	7	3020
## 99	2002	Month	7	3669
## 100	2003	Month	7	4431
## 101	2004	Month	7	4831
## 102	2005	Month	7	5597
## 103	2006	Month	7	5974
## 104	2007	Month	7	6155
## 105	2008	Month	7	7059
## 106	2009	Month	7	7752
## 107	2010	Month	7	8202
## 108	2011	Month	7	8844
## 109	2012	Month	7	9596
## 110	2013	Month	7	9846
## 111	2014	Month	7	10843
## 112	2015	Month	7	11545
## 113	2000	Month	8	2696
## 114	2001	Month	8	3432
## 115	2002	Month	8	3998
## 116	2003	Month	8	4877
## 117	2004	Month	8	5220
## 118	2005	Month	8	5943
## 119	2006	Month	8	6260
## 120	2007	Month	8	6444
## 121	2008	Month	8	7267
## 122	2009	Month	8	8186
## 123	2010	Month	8	8482
## 124	2011	Month	8	9103
## 125	2012	Month	8	10028
## 126	2013	Month	8	10202
## 127	2014	Month	8	11062
## 128	2015	Month	8	11777
## 129	2000	Month	9	2255
## 130	2001	Month	9	3042
## 131	2002	Month	9	3574
## 132	2003	Month	9	4390
## 133	2004	Month	9	4836
## 134	2005	Month	9	5628
## 135	2006	Month	9	5774
## 136	2007	Month	9	6022
## 137	2008	Month	9	6862
## 138	2009	Month	9	7844
## 139	2010	Month	9	8122
## 140	2011	Month	9	8606
## 141	2012	Month	9	9549
## 142	2013	Month	9	9764
## 143	2014	Month	9	10653
## 144	2015	Month	9	11405

## 145	2000	Month	10	2022
## 146	2001	Month	10	2825
## 147	2002	Month	10	3344
## 148	2003	Month	10	4181
## 149	2004	Month	10	4627
## 150	2005	Month	10	5419
## 151	2006	Month	10	5542
## 152	2007	Month	10	5813
## 153	2008	Month	10	6643
## 154	2009	Month	10	7596
## 155	2010	Month	10	7891
## 156	2011	Month	10	8392
## 157	2012	Month	10	9313
## 158	2013	Month	10	9538
## 159	2014	Month	10	10452
## 160	2015	Month	10	11168
## 161	2000	Month	11	1817
## 162	2001	Month	11	2598
## 163	2002	Month	11	3112
## 164	2003	Month	11	3946
## 165	2004	Month	11	4414
## 166	2005	Month	11	5209
## 167	2006	Month	11	5309
## 168	2007	Month	11	5582
## 169	2008	Month	11	6441
## 170	2009	Month	11	7351
## 171	2010	Month	11	7677
## 172	2011	Month	11	8151
## 173	2012	Month	11	9091
## 174	2013	Month	11	9306
## 175	2014	Month	11	10224
## 176	2015	Month	11	10922
## 177	2000	Month	12	1594
## 178	2001	Month	12	2368
## 179	2002	Month	12	2908
## 180	2003	Month	12	3706
## 181	2004	Month	12	4191
## 182	2005	Month	12	4984
## 183	2006	Month	12	5073
## 184	2007	Month	12	5342
## 185	2008	Month	12	6205
## 186	2009	Month	12	7121
## 187	2010	Month	12	7435
## 188	2011	Month	12	7929
## 189	2012	Month	12	8865
## 190	2013	Month	12	9058
## 191	2014	Month	12	10011
## 192	2015	Month	12	10721

5. Use the unite function on the data from part 4 to merge “Time” and “Time_Interval” create a new column to read as “1_Month”

```
#unite( data, name_of_new_column, names_of_col_merging_[LIST] , sep = " ", remove = TRUE)
new_table_unite=unite(new_table_separate, "New_Time_Month", c("Time_Interval","Time"), sep = "_",remove=TRUE)
new_table_unite
```

##	Year	New_Time_Month	Revenue
## 1	2000	1_Month	1008
## 2	2001	1_Month	1472
## 3	2002	1_Month	2133
## 4	2003	1_Month	2844
## 5	2004	1_Month	3453
## 6	2005	1_Month	4034
## 7	2006	1_Month	4548
## 8	2007	1_Month	5039
## 9	2008	1_Month	5599
## 10	2009	1_Month	6217
## 11	2010	1_Month	6834
## 12	2011	1_Month	7357
## 13	2012	1_Month	7991
## 14	2013	1_Month	8695
## 15	2014	1_Month	9305
## 16	2015	1_Month	9950
## 17	2000	2_Month	1119
## 18	2001	2_Month	1657
## 19	2002	2_Month	2307
## 20	2003	2_Month	3011
## 21	2004	2_Month	3668
## 22	2005	2_Month	4201
## 23	2006	2_Month	4646
## 24	2007	2_Month	5165
## 25	2008	2_Month	5788
## 26	2009	2_Month	6372
## 27	2010	2_Month	6965
## 28	2011	2_Month	7464
## 29	2012	2_Month	8119
## 30	2013	2_Month	8816
## 31	2014	2_Month	9437
## 32	2015	2_Month	10186
## 33	2000	3_Month	1204
## 34	2001	3_Month	1844
## 35	2002	3_Month	2510
## 36	2003	3_Month	3155
## 37	2004	3_Month	3817
## 38	2005	3_Month	4276
## 39	2006	3_Month	4733
## 40	2007	3_Month	5367
## 41	2008	3_Month	5919
## 42	2009	3_Month	6596
## 43	2010	3_Month	7054
## 44	2011	3_Month	7695
## 45	2012	3_Month	8291
## 46	2013	3_Month	9027
## 47	2014	3_Month	9604
## 48	2015	3_Month	10361
## 49	2000	4_Month	1292
## 50	2001	4_Month	1962
## 51	2002	4_Month	2668
## 52	2003	4_Month	3286
## 53	2004	4_Month	3965

## 54	2005	4_Month	4392
## 55	2006	4_Month	4817
## 56	2007	4_Month	5494
## 57	2008	4_Month	6107
## 58	2009	4_Month	6702
## 59	2010	4_Month	7263
## 60	2011	4_Month	7874
## 61	2012	4_Month	8506
## 62	2013	4_Month	9140
## 63	2014	4_Month	9829
## 64	2015	4_Month	10552
## 65	2000	5_Month	1560
## 66	2001	5_Month	2347
## 67	2002	5_Month	2909
## 68	2003	5_Month	3739
## 69	2004	5_Month	4176
## 70	2005	5_Month	4743
## 71	2006	5_Month	5163
## 72	2007	5_Month	5701
## 73	2008	5_Month	6360
## 74	2009	5_Month	7077
## 75	2010	5_Month	7721
## 76	2011	5_Month	8155
## 77	2012	5_Month	8919
## 78	2013	5_Month	9343
## 79	2014	5_Month	10259
## 80	2015	5_Month	10863
## 81	2000	6_Month	2007
## 82	2001	6_Month	2640
## 83	2002	6_Month	3279
## 84	2003	6_Month	4024
## 85	2004	6_Month	4593
## 86	2005	6_Month	5186
## 87	2006	6_Month	5635
## 88	2007	6_Month	5925
## 89	2008	6_Month	6579
## 90	2009	6_Month	7305
## 91	2010	6_Month	7989
## 92	2011	6_Month	8634
## 93	2012	6_Month	9222
## 94	2013	6_Month	9550
## 95	2014	6_Month	10638
## 96	2015	6_Month	11115
## 97	2000	7_Month	2496
## 98	2001	7_Month	3020
## 99	2002	7_Month	3669
## 100	2003	7_Month	4431
## 101	2004	7_Month	4831
## 102	2005	7_Month	5597
## 103	2006	7_Month	5974
## 104	2007	7_Month	6155
## 105	2008	7_Month	7059
## 106	2009	7_Month	7752
## 107	2010	7_Month	8202

## 108 2011	7_Month	8844
## 109 2012	7_Month	9596
## 110 2013	7_Month	9846
## 111 2014	7_Month	10843
## 112 2015	7_Month	11545
## 113 2000	8_Month	2696
## 114 2001	8_Month	3432
## 115 2002	8_Month	3998
## 116 2003	8_Month	4877
## 117 2004	8_Month	5220
## 118 2005	8_Month	5943
## 119 2006	8_Month	6260
## 120 2007	8_Month	6444
## 121 2008	8_Month	7267
## 122 2009	8_Month	8186
## 123 2010	8_Month	8482
## 124 2011	8_Month	9103
## 125 2012	8_Month	10028
## 126 2013	8_Month	10202
## 127 2014	8_Month	11062
## 128 2015	8_Month	11777
## 129 2000	9_Month	2255
## 130 2001	9_Month	3042
## 131 2002	9_Month	3574
## 132 2003	9_Month	4390
## 133 2004	9_Month	4836
## 134 2005	9_Month	5628
## 135 2006	9_Month	5774
## 136 2007	9_Month	6022
## 137 2008	9_Month	6862
## 138 2009	9_Month	7844
## 139 2010	9_Month	8122
## 140 2011	9_Month	8606
## 141 2012	9_Month	9549
## 142 2013	9_Month	9764
## 143 2014	9_Month	10653
## 144 2015	9_Month	11405
## 145 2000	10_Month	2022
## 146 2001	10_Month	2825
## 147 2002	10_Month	3344
## 148 2003	10_Month	4181
## 149 2004	10_Month	4627
## 150 2005	10_Month	5419
## 151 2006	10_Month	5542
## 152 2007	10_Month	5813
## 153 2008	10_Month	6643
## 154 2009	10_Month	7596
## 155 2010	10_Month	7891
## 156 2011	10_Month	8392
## 157 2012	10_Month	9313
## 158 2013	10_Month	9538
## 159 2014	10_Month	10452
## 160 2015	10_Month	11168
## 161 2000	11_Month	1817


```
## 162 2001      11_Month    2598
## 163 2002      11_Month    3112
## 164 2003      11_Month    3946
## 165 2004      11_Month    4414
## 166 2005      11_Month    5209
## 167 2006      11_Month    5309
## 168 2007      11_Month    5582
## 169 2008      11_Month    6441
## 170 2009      11_Month    7351
## 171 2010      11_Month    7677
## 172 2011      11_Month    8151
## 173 2012      11_Month    9091
## 174 2013      11_Month    9306
## 175 2014      11_Month   10224
## 176 2015      11_Month   10922
## 177 2000      12_Month    1594
## 178 2001      12_Month    2368
## 179 2002      12_Month    2908
## 180 2003      12_Month    3706
## 181 2004      12_Month    4191
## 182 2005      12_Month    4984
## 183 2006      12_Month    5073
## 184 2007      12_Month    5342
## 185 2008      12_Month    6205
## 186 2009      12_Month    7121
## 187 2010      12_Month    7435
## 188 2011      12_Month    7929
## 189 2012      12_Month    8865
## 190 2013      12_Month    9058
## 191 2014      12_Month   10011
## 192 2015      12_Month   10721
```

6. Use the spread function on the data from part 5 such that the columns are the years and keep the month in a column.

```
#spread( data, key: columns_to_spread, value: values_of_spreadColumns, fill= NA, convert = FALSE)
spread(new_table_unite,Year,Revenue)
```

```
##      New_Time_Month 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 1      10_Month 2022 2825 3344 4181 4627 5419 5542 5813 6643 7596 7891
## 2      11_Month 1817 2598 3112 3946 4414 5209 5309 5582 6441 7351 7677
## 3      12_Month 1594 2368 2908 3706 4191 4984 5073 5342 6205 7121 7435
## 4        1_Month 1008 1472 2133 2844 3453 4034 4548 5039 5599 6217 6834
## 5        2_Month 1119 1657 2307 3011 3668 4201 4646 5165 5788 6372 6965
## 6        3_Month 1204 1844 2510 3155 3817 4276 4733 5367 5919 6596 7054
## 7        4_Month 1292 1962 2668 3286 3965 4392 4817 5494 6107 6702 7263
## 8        5_Month 1560 2347 2909 3739 4176 4743 5163 5701 6360 7077 7721
## 9        6_Month 2007 2640 3279 4024 4593 5186 5635 5925 6579 7305 7989
## 10       7_Month 2496 3020 3669 4431 4831 5597 5974 6155 7059 7752 8202
## 11       8_Month 2696 3432 3998 4877 5220 5943 6260 6444 7267 8186 8482
## 12       9_Month 2255 3042 3574 4390 4836 5628 5774 6022 6862 7844 8122
##      2011 2012 2013 2014 2015
## 1  8392 9313 9538 10452 11168
## 2  8151 9091 9306 10224 10922
## 3  7929 8865 9058 10011 10721
```

##	4	7357	7991	8695	9305	9950
##	5	7464	8119	8816	9437	10186
##	6	7695	8291	9027	9604	10361
##	7	7874	8506	9140	9829	10552
##	8	8155	8919	9343	10259	10863
##	9	8634	9222	9550	10638	11115
##	10	8844	9596	9846	10843	11545
##	11	9103	10028	10202	11062	11777
##	12	8606	9549	9764	10653	11405