

(STAT 745) Individual Project: EDA

Overview

This report will explore the dataset to support the creation of a predictive model.

```
library(here)
library(magrittr)
library(dplyr)
library(skimr)
```

Loading the Data

Load train and test subsets and combine into data frames with a 'Subset' label.

```
# Load data subsets.
load(here("Data.train.RData"))
load(here("Data.test.RData"))

# Combine subsets.
Data.df <-
  dplyr::bind_rows(
    list(TRAIN = Data.train,
         TEST = Data.test),
    .id = "Subset")
```

```
head(Data.df)
```

	Subset	Y	X1	X2	X3	X4	X5	X6	X7	X8
1	TRAIN	2651.847	190.2752	143.97652	151.7833	186.4703	156.51065	148.6427	171.89960	197.443
2	TRAIN	11138.524	187.3541	163.27790	181.8418	172.3075	173.42362	165.2064	191.49909	184.486
3	TRAIN	5243.665	121.5722	79.75382	99.0256	117.0735	97.87589	122.1271	99.76079	121.512
4	TRAIN	2913.020	165.4992	129.64572	141.8949	175.6651	140.55434	120.1036	130.93532	148.308
5	TRAIN	1274.079	156.8711	142.60402	150.7077	159.6849	129.32182	120.1012	139.95636	160.943
6	TRAIN	7093.835	232.3824	197.11017	220.6493	225.4648	199.04439	196.5903	210.19434	202.973
	X12	X13	X14	X15	X16	X17	X18	X19	X20	X21
1	189.4911	170.9808	187.0820	172.8512	178.9792	170.3895	190.3821	154.33964	180.9958	158.1956
2	200.8819	180.8596	195.2746	179.7441	167.2038	189.8296	183.8584	151.59114	171.6527	161.7206
3	158.7075	122.8142	110.3308	141.2242	117.3978	113.2675	103.3438	98.87883	124.5750	125.0871
4	163.9969	137.4046	143.4139	152.1899	143.0091	146.3000	138.2800	129.95202	158.5156	138.6803
5	165.6185	139.4516	163.9141	155.6802	148.8496	143.8891	167.8937	148.78064	173.1553	152.2767
6	223.7867	209.5524	237.6196	227.2100	221.6235	222.8292	232.0893	212.20279	228.7125	210.2681
	X26	X27	X28	X29	X30	X31	X32	X33	X34	X35
1	189.9048	183.8973	196.2291	199.1824	159.8516	160.1199	164.9472	175.5786	157.8679	204.5142
2	176.6973	171.4377	169.2847	191.4913	185.2922	166.0120	183.5489	165.8716	169.4822	193.3774

3	122.1213	121.3448	123.2495	147.3061	112.2195	110.0080	126.6346	112.0581	147.5703	134.2955
4	165.3876	139.9958	145.8997	176.8595	138.7055	151.4801	138.2511	159.5250	140.0390	153.7903
5	156.7271	154.3948	168.6283	186.5658	133.0860	139.9109	155.8428	162.2954	140.5824	156.9258
6	211.1699	202.1341	207.4115	246.5825	221.0299	181.7204	226.0357	202.8005	229.3678	225.2104
	X40	X41	X42	X43	X44	X45	X46	X47	X48	X49
1	165.9454	161.0893	186.5573	181.0184	168.4939	155.2141	165.2016	173.7880	162.1056	204.0403
2	149.8757	169.7056	167.7227	189.1325	167.6121	158.0837	169.3665	182.6245	176.9324	187.7063
3	107.9653	119.0335	84.0098	114.2309	112.5771	115.2432	103.0133	142.4198	101.8722	140.4499
4	152.1070	143.1821	154.5452	161.9060	130.1764	145.6968	142.3990	145.2875	129.0462	156.7216
5	170.4956	153.2533	146.9815	157.4408	154.7867	160.3062	128.9857	154.8565	125.2632	149.7415
6	218.5095	217.5136	216.8779	217.5554	210.8703	214.5118	211.0018	238.0285	186.1718	209.2742
	X54	X55	X56	X57	X58	X59	X60	X61	X62	X63
1	161.1614	156.2437	195.6223	175.1965	195.8409	159.0638	164.5755	167.7611	145.7529	158.0565
2	166.5849	172.4211	184.2249	188.4637	181.6099	169.3594	150.7625	187.3125	181.8074	169.2095
3	118.2258	123.6946	105.3679	129.5990	149.2307	119.2903	112.7646	127.7381	125.7804	81.7235
4	118.2854	149.1520	185.1071	150.1289	145.9323	136.2323	133.1889	141.4037	164.7227	147.8011
5	129.2905	167.5671	163.6706	142.0493	153.0318	141.2782	154.5048	150.0959	150.1901	141.6830
6	220.8697	208.2375	233.1126	246.7481	229.1053	205.7242	221.0139	227.6108	215.1067	183.0254
	X68	X69	X70	X71	X72	X73	X74	X75	X76	X77
1	153.02558	147.7750	149.46138	183.1023	157.5490	176.7347	153.2959	194.1464	194.4869	175.553
2	164.22532	167.0048	141.58409	203.2107	164.9747	168.5169	163.9777	184.4621	185.4475	166.937
3	80.22515	120.6989	61.05903	126.8386	104.4448	110.5760	103.5457	144.9413	135.9236	133.140
4	123.14236	128.6165	148.80208	145.5930	140.2170	156.5556	130.5368	147.1809	135.4346	155.211
5	133.92236	145.3575	141.97908	138.2948	140.9657	145.2275	157.9091	160.8787	154.5288	154.240
6	202.74809	203.7470	209.94033	220.2130	221.4984	218.4386	194.2187	220.3678	242.8217	197.823
	X81	X82	X83	X84	X85	X86	X87	X88	X89	X90
1	173.6581	179.4466	169.8282	165.8769	172.0811	166.1395	167.2687	153.0261	193.3265	188.2962
2	179.3780	170.1728	192.6931	187.2476	191.4529	156.8783	184.8603	148.0441	186.9830	196.3248
3	112.9946	126.9060	127.6400	142.7316	126.4520	101.3542	114.0817	124.5336	103.8253	132.7848
4	152.5364	138.3971	142.7609	156.7981	159.1366	144.8108	161.3606	142.2514	153.0177	158.6251
5	146.8272	173.1438	163.4618	160.9412	153.1834	124.8014	149.0323	156.9803	160.8083	161.4221
6	200.3824	207.7934	235.9465	243.9284	240.1148	193.4679	217.8605	195.7473	248.7948	232.5632
	X95	X96	X97	X98	X99	X100	X101	X102	X103	X104
1	161.3741	150.5578	201.7703	151.6335	183.8449	187.8406	182.9719	176.5657	174.4933	196.4328
2	188.0247	147.8279	188.3031	186.3663	179.4196	194.3465	173.0327	155.7836	186.8129	166.5490
3	132.1839	131.0429	137.0399	131.4063	113.4326	111.3325	116.7455	128.7372	112.0441	103.0925
4	139.3936	150.5044	177.0203	137.9799	154.3753	149.2236	144.5166	138.4869	162.6022	168.4266
5	164.4596	156.4609	158.9963	138.2733	144.3009	157.7780	155.8910	158.7553	167.6108	169.9792
6	250.2296	228.8293	234.2316	212.0430	193.8184	235.9374	212.6803	234.2633	232.1351	236.6554
	X109	X110	X111	X112	X113	X114	X115	X116	X117	X118
1	188.0414	170.4011	198.4534	159.7283	157.3595	171.8311	180.5850	163.0439	175.2278	159.1415
2	188.0961	167.5971	190.7035	173.3698	164.5685	172.7110	178.7160	163.7623	174.0169	161.0934
3	126.4743	123.1635	120.6548	116.1058	100.2987	116.5577	138.4449	103.2853	91.4304	102.7650
4	146.5788	144.4557	157.5633	145.6816	119.1198	153.6854	173.9040	138.1342	137.9500	144.8583
5	153.0245	150.8929	170.4957	139.1781	145.7941	157.5177	178.0832	138.7221	153.2255	146.1542
6	232.4422	212.4368	226.7012	208.7807	209.6966	232.9099	235.6005	235.6735	224.3841	210.1687
	X122	X123	X124	X125	X126	X127	X128	X129	X130	X131
1	173.1394	176.2271	170.5302	179.9718	171.9809	185.4365	155.77938	155.8573	198.1534	184.6739
2	175.1476	191.8524	168.9681	175.4244	178.3329	180.7142	195.40879	157.5818	176.1117	162.2730
3	114.3544	132.8511	103.1675	116.6601	144.9051	117.7660	94.43294	115.9584	120.8620	120.4363
4	157.9706	157.8891	158.0380	158.3967	178.7961	145.2719	141.41390	152.1806	152.1110	162.5889

```

5 170.6687 148.6311 157.1162 173.3105 170.1513 153.6995 139.44773 130.7491 150.8242 177.6694
6 226.1086 224.6901 206.0081 242.4243 218.7541 217.6847 203.76055 200.5651 227.4442 212.7767
      X136      X137      X138      X139      X140
1 208.2443 162.1775 181.8940 165.6101 188.0576
2 194.9840 173.5201 197.7736 146.5603 177.1221
3 127.8460  96.5828 155.7711 122.1913 135.8853
4 165.0938 143.9547 150.8131 133.8978 150.1762
5 171.7028 133.5584 164.6517 139.4064 163.1669
6 240.9290 220.9919 230.5252 203.4847 229.1268

```

Exploring the Data

Summary of all data points:

```
Data.df %>% skim_without_charts() %>% print()
```

```
-- Data Summary -----
```

Name	Values
Number of rows	Piped data
Number of columns	10400
	142

```
-----
Column type frequency:
```

character	1
numeric	141

```
-----
Group variables      None
```

```
-- Variable type: character -----
```

```
# A tibble: 1 x 8
```

	skim_variable	n_missing	complete_rate	min	max	empty	n_unique	whitespace
*	<chr>	<int>	<dbl>	<int>	<int>	<int>	<int>	<int>
1	Subset	0	1	4	5	0	2	0

```
-- Variable type: numeric -----
```

```
# A tibble: 141 x 10
```

	skim_variable	n_missing	complete_rate	mean	sd	p0	p25	p50	p75	p100
*	<chr>	<int>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>
1	Y	5200	0.5	9234.	5116.	-7028.	5574.	9148.	12727.	31402.
2	X1	0	1	180.	23.3	73.5	165.	181.	196.	268.
3	X2	0	1	166.	26.7	72.3	148.	167.	185.	263.
4	X3	0	1	178.	25.6	80.5	160.	179.	196.	262.
5	X4	0	1	179.	24.6	74.4	163.	179.	196.	264.
6	X5	0	1	167.	24.8	76.1	150.	167.	184.	247.
7	X6	0	1	156.	25.0	54.7	139.	157.	174.	240.
8	X7	0	1	171.	25.9	72.8	153.	172.	189.	252.
9	X8	0	1	178.	23.8	89.3	162.	178.	194.	257.
10	X9	0	1	163.	26.0	65.6	145.	165.	182.	247.
11	X10	0	1	169.	23.2	79.0	153.	169.	184.	250.
12	X11	0	1	161.	25.1	66.0	144.	161.	179.	242.
13	X12	0	1	197.	27.4	83.2	179.	198.	216.	290.

14	X13	0	1	171.	24.8	81.0	154.	172.	188.	256.
15	X14	0	1	190.	30.6	75.0	170.	193.	213.	283.
16	X15	0	1	181.	26.9	74.1	162.	182.	200.	272.
17	X16	0	1	174.	25.0	56.8	157.	175.	191.	262.
18	X17	0	1	183.	26.3	84.3	165.	183.	201.	272.
19	X18	0	1	190.	28.0	77.2	171.	191.	210.	276.
20	X19	0	1	176.	26.7	66.8	158.	178.	195.	260.
21	X20	0	1	177.	25.6	84.9	159.	177.	195.	262.
22	X21	0	1	167.	22.4	67.4	152.	167.	183.	242.
23	X22	0	1	186.	25.7	86.6	168.	186.	204.	280.
24	X23	0	1	176.	25.8	64.4	158.	176.	194.	261.
25	X24	0	1	173.	23.7	78.6	157.	173.	189.	255.
26	X25	0	1	176.	25.0	81.9	160.	177.	194.	261.
27	X26	0	1	184.	24.6	95.1	167.	184.	201.	267.
28	X27	0	1	173.	24.5	78.6	156.	174.	190.	255.
29	X28	0	1	172.	23.6	76.7	156.	173.	188.	256.
30	X29	0	1	199.	27.8	103.	181.	200.	218.	287.
31	X30	0	1	175.	27.4	68.8	156.	176.	194.	255.
32	X31	0	1	176.	26.6	78.2	158.	178.	195.	264.
33	X32	0	1	179.	26.5	82.7	161.	180.	198.	278.
34	X33	0	1	176.	23.3	79.2	160.	176.	192.	255.
35	X34	0	1	171.	24.9	83.1	154.	171.	188.	253.
36	X35	0	1	190.	28.2	89.3	171.	191.	211.	281.
37	X36	0	1	172.	25.7	65.3	154.	173.	190.	262.
38	X37	0	1	170.	22.3	80.8	155.	170.	185.	250.
39	X38	0	1	182.	26.3	86.7	163.	182.	200.	269.
40	X39	0	1	167.	24.1	74.9	151.	168.	184.	253.
41	X40	0	1	179.	26.9	84.4	162.	181.	198.	278.
42	X41	0	1	167.	24.8	65.8	150.	167.	185.	254.
43	X42	0	1	178.	26.8	62.3	160.	180.	197.	258.
44	X43	0	1	180.	26.9	75.0	162.	181.	199.	271.
45	X44	0	1	176.	26.8	58.8	158.	177.	195.	268.
46	X45	0	1	179.	24.3	78.4	163.	179.	195.	268.
47	X46	0	1	171.	25.4	75.4	154.	172.	189.	253.
48	X47	0	1	195.	26.8	70.4	177.	196.	214.	279.
49	X48	0	1	161.	25.5	55.5	144.	162.	179.	250.
50	X49	0	1	180.	25.2	69.3	163.	180.	197.	261.
51	X50	0	1	193.	26.2	82.4	175.	194.	212.	290.
52	X51	0	1	175.	24.1	82.6	158.	175.	191.	261.
53	X52	0	1	177.	25.3	84.9	160.	177.	194.	272.
54	X53	0	1	188.	25.2	89.1	171.	189.	205.	272.
55	X54	0	1	164.	24.0	63.3	147.	164.	180.	238.
56	X55	0	1	174.	24.0	85.5	158.	175.	191.	253.
57	X56	0	1	190.	26.5	90.6	172.	191.	209.	292.
58	X57	0	1	175.	25.1	78.7	158.	175.	193.	268.
59	X58	0	1	184.	25.5	86.8	167.	185.	202.	268.
60	X59	0	1	168.	23.9	74.3	151.	168.	185.	248.
61	X60	0	1	175.	27.2	69.5	156.	175.	194.	266.
62	X61	0	1	178.	26.0	70.1	160.	178.	196.	259.
63	X62	0	1	176.	26.0	84.6	158.	176.	195.	260.
64	X63	0	1	163.	23.7	79.0	147.	164.	180.	246.

65	X64	0	1	177.	24.3	62.4	161.	178.	194.	255.
66	X65	0	1	176.	24.0	86.3	160.	177.	193.	258.
67	X66	0	1	177.	24.6	91.9	160.	178.	195.	275.
68	X67	0	1	179.	26.3	80.5	161.	180.	198.	266.
69	X68	0	1	160.	25.8	69.1	142.	161.	178.	243.
70	X69	0	1	159.	22.7	64.9	143.	159.	175.	238.
71	X70	0	1	167.	27.2	61.1	148.	169.	186.	248.
72	X71	0	1	178.	26.0	82.3	160.	179.	196.	267.
73	X72	0	1	165.	22.2	80.7	150.	165.	180.	252.
74	X73	0	1	177.	24.6	84.9	160.	177.	194.	267.
75	X74	0	1	166.	24.8	68.4	149.	167.	184.	250.
76	X75	0	1	183.	23.4	69.4	167.	183.	199.	274.
77	X76	0	1	180.	26.5	79.0	162.	181.	199.	276.
78	X77	0	1	176.	25.1	80.4	159.	177.	193.	258.
79	X78	0	1	166.	24.1	67.7	149.	166.	182.	253.
80	X79	0	1	182.	26.1	82.1	164.	182.	200.	277.
81	X80	0	1	181.	26.5	83.7	163.	182.	200.	270.
82	X81	0	1	172.	27.3	71.8	153.	173.	191.	260.
83	X82	0	1	184.	25.6	92.3	166.	184.	202.	272.
84	X83	0	1	184.	30.2	68.7	163.	185.	205.	280.
85	X84	0	1	184.	27.2	73.9	166.	185.	203.	272.
86	X85	0	1	188.	27.7	90.7	169.	189.	208.	271.
87	X86	0	1	166.	23.1	79.5	150.	167.	181.	237.
88	X87	0	1	172.	26.9	77.6	154.	173.	191.	254.
89	X88	0	1	168.	24.1	81.3	153.	169.	185.	247.
90	X89	0	1	180.	25.2	89.4	162.	180.	197.	267.
91	X90	0	1	186.	26.4	79.8	168.	186.	204.	277.
92	X91	0	1	167.	27.1	61.6	148.	168.	186.	257.
93	X92	0	1	172.	22.2	81.4	157.	172.	187.	248.
94	X93	0	1	194.	26.6	97.2	176.	195.	213.	280.
95	X94	0	1	177.	25.8	89.4	159.	177.	195.	266.
96	X95	0	1	182.	31.2	73.0	160.	184.	204.	273.
97	X96	0	1	173.	23.1	87.0	157.	173.	189.	269.
98	X97	0	1	196.	26.2	92.2	178.	196.	214.	290.
99	X98	0	1	174.	27.2	77.7	155.	175.	193.	263.
100	X99	0	1	168.	23.6	86.6	152.	168.	184.	249.
101	X100	0	1	188.	28.6	72.8	168.	188.	208.	279.
102	X101	0	1	181.	25.1	60.3	164.	182.	198.	272.
103	X102	0	1	181.	26.9	82.0	163.	182.	201.	269.
104	X103	0	1	190.	26.6	72.6	172.	190.	208.	277.
105	X104	0	1	187.	26.6	89.1	169.	187.	206.	273.
106	X105	0	1	169.	27.4	69.0	151.	171.	189.	250.
107	X106	0	1	184.	24.7	87.5	167.	184.	201.	279.
108	X107	0	1	167.	24.6	74.0	149.	167.	184.	253.
109	X108	0	1	160.	24.0	68.5	144.	161.	177.	232.
110	X109	0	1	177.	25.6	78.1	160.	178.	195.	260.
111	X110	0	1	176.	25.1	84.3	160.	177.	193.	256.
112	X111	0	1	190.	24.8	99.5	173.	190.	207.	268.
113	X112	0	1	170.	25.7	77.1	153.	171.	188.	275.
114	X113	0	1	163.	24.2	76.4	147.	164.	180.	247.
115	X114	0	1	176.	25.3	65.6	159.	176.	194.	266.

116	X115	0	1	197.	27.3	77.6	178.	198.	216.	291.
117	X116	0	1	172.	26.2	65.3	154.	173.	190.	258.
118	X117	0	1	167.	23.5	63.9	151.	167.	183.	250.
119	X118	0	1	172.	24.2	71.6	156.	173.	189.	249.
120	X119	0	1	170.	22.9	81.2	155.	171.	186.	255.
121	X120	0	1	184.	28.9	85.2	164.	185.	205.	273.
122	X121	0	1	170.	25.3	77.5	153.	170.	188.	260.
123	X122	0	1	184.	24.1	93.8	167.	184.	200.	271.
124	X123	0	1	180.	26.1	78.0	162.	181.	199.	264.
125	X124	0	1	176.	23.6	74.7	160.	176.	192.	262.
126	X125	0	1	186.	27.0	78.9	167.	187.	205.	268.
127	X126	0	1	185.	24.0	92.9	168.	185.	201.	267.
128	X127	0	1	175.	25.1	74.0	158.	176.	193.	258.
129	X128	0	1	167.	25.4	67.0	150.	168.	185.	248.
130	X129	0	1	172.	25.7	75.2	154.	172.	190.	261.
131	X130	0	1	181.	25.3	87.3	164.	181.	199.	265.
132	X131	0	1	178.	22.7	87.4	162.	178.	193.	263.
133	X132	0	1	167.	24.8	46.8	150.	167.	184.	255.
134	X133	0	1	190.	27.5	86.6	171.	191.	209.	281.
135	X134	0	1	179.	25.7	85.4	162.	180.	197.	265.
136	X135	0	1	186.	26.2	94.3	168.	186.	204.	276.
137	X136	0	1	190.	27.0	79.0	172.	191.	209.	286.
138	X137	0	1	163.	23.1	54.6	148.	164.	179.	241.
139	X138	0	1	193.	28.2	91.2	174.	195.	213.	285.
140	X139	0	1	169.	24.2	78.6	152.	169.	186.	252.
141	X140	0	1	182.	26.5	72.5	164.	183.	200.	265.

Summary of data grouped by subsets:

```
Data.df %>% group_by(Subset) %>% skim_without_charts() %>% print()
```

```
-- Data Summary -----
```

	Values
Name	Piped data
Number of rows	10400
Number of columns	142

```
-----
```

Column type frequency:	
numeric	141

```
-----
```

Group variables	Subset
-----------------	--------

```
-- Variable type: numeric -----
```

```
# A tibble: 282 x 11
```

	skim_variable	Subset	n_missing	complete_rate	mean	sd	p0	p25	p50	p75	p100
*	<chr>	<chr>	<int>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>
1	Y	TEST	5200	0	NaN	NA	NA	NA	NA	NA	NA
2	Y	TRAIN	0	1	9234.	5116.	-7028.	5574.	9148.	12727.	31416.
3	X1	TEST	0	1	181.	23.1	91.1	165.	181.	196.	211.
4	X1	TRAIN	0	1	180.	23.5	73.5	164.	180.	196.	211.
5	X2	TEST	0	1	167.	26.4	87.0	149.	167.	185.	203.
6	X2	TRAIN	0	1	166.	27.0	72.3	147.	166.	185.	203.

7	X3	TEST	0	1	179.	25.7	80.5	161.	179.	197.	2
8	X3	TRAIN	0	1	178.	25.6	83.2	160.	179.	196.	2
9	X4	TEST	0	1	180.	24.4	94.8	163.	180.	197.	2
10	X4	TRAIN	0	1	179.	24.8	74.4	162.	179.	196.	2
11	X5	TEST	0	1	167.	24.6	76.1	151.	167.	184.	2
12	X5	TRAIN	0	1	166.	25.0	80.3	149.	166.	183.	2
13	X6	TEST	0	1	157.	24.9	64.4	140.	158.	174.	2
14	X6	TRAIN	0	1	156.	25.1	54.7	139.	157.	174.	2
15	X7	TEST	0	1	171.	25.5	73.0	154.	172.	189.	2
16	X7	TRAIN	0	1	171.	26.4	72.8	153.	172.	189.	2
17	X8	TEST	0	1	178.	23.7	89.3	162.	178.	194.	2
18	X8	TRAIN	0	1	178.	23.9	96.2	162.	178.	194.	2
19	X9	TEST	0	1	164.	26.0	65.6	146.	165.	182.	2
20	X9	TRAIN	0	1	163.	26.0	78.3	145.	164.	181.	2
21	X10	TEST	0	1	169.	23.3	79.0	154.	169.	185.	2
22	X10	TRAIN	0	1	168.	23.2	88.3	153.	169.	184.	2
23	X11	TEST	0	1	162.	25.1	72.4	144.	161.	179.	2
24	X11	TRAIN	0	1	161.	25.1	66.0	144.	161.	178.	2
25	X12	TEST	0	1	198.	27.3	83.2	179.	199.	216.	2
26	X12	TRAIN	0	1	197.	27.4	95.9	179.	198.	216.	2
27	X13	TEST	0	1	171.	24.5	81.6	155.	172.	188.	2
28	X13	TRAIN	0	1	170.	25.1	81.0	153.	171.	188.	2
29	X14	TEST	0	1	191.	30.2	89.1	171.	193.	213.	2
30	X14	TRAIN	0	1	190.	30.9	75.0	169.	193.	213.	2
31	X15	TEST	0	1	181.	26.9	74.1	163.	183.	200.	2
32	X15	TRAIN	0	1	181.	26.9	81.4	162.	182.	200.	2
33	X16	TEST	0	1	174.	24.8	56.8	158.	175.	192.	2
34	X16	TRAIN	0	1	174.	25.2	73.7	157.	175.	191.	2
35	X17	TEST	0	1	183.	26.1	92.7	166.	184.	201.	2
36	X17	TRAIN	0	1	183.	26.5	84.3	165.	183.	201.	2
37	X18	TEST	0	1	191.	27.6	86.0	172.	191.	210.	2
38	X18	TRAIN	0	1	190.	28.3	77.2	170.	191.	210.	2
39	X19	TEST	0	1	176.	26.6	66.8	159.	179.	195.	2
40	X19	TRAIN	0	1	175.	26.7	78.6	158.	178.	194.	2
41	X20	TEST	0	1	177.	25.4	84.9	160.	178.	195.	2
42	X20	TRAIN	0	1	176.	25.9	88.1	159.	177.	195.	2
43	X21	TEST	0	1	167.	22.4	85.4	152.	167.	183.	2
44	X21	TRAIN	0	1	167.	22.4	67.4	152.	167.	182.	2
45	X22	TEST	0	1	186.	25.7	86.6	169.	187.	204.	2
46	X22	TRAIN	0	1	185.	25.7	93.0	167.	186.	203.	2
47	X23	TEST	0	1	176.	25.6	86.4	158.	177.	194.	2
48	X23	TRAIN	0	1	175.	25.9	64.4	157.	176.	194.	2
49	X24	TEST	0	1	173.	23.5	89.0	158.	174.	190.	2
50	X24	TRAIN	0	1	173.	23.8	78.6	156.	173.	189.	2
51	X25	TEST	0	1	177.	24.7	87.2	161.	177.	194.	2
52	X25	TRAIN	0	1	176.	25.2	81.9	159.	177.	193.	2
53	X26	TEST	0	1	184.	24.4	100.	167.	185.	201.	2
54	X26	TRAIN	0	1	184.	24.8	95.1	167.	184.	201.	2
55	X27	TEST	0	1	173.	24.4	83.8	157.	174.	190.	2
56	X27	TRAIN	0	1	172.	24.6	78.6	156.	173.	190.	2
57	X28	TEST	0	1	173.	23.6	85.1	157.	173.	188.	2

58	X28	TRAIN	0	1	172.	23.6	76.7	156.	173.	188.	2
59	X29	TEST	0	1	199.	27.6	103.	181.	200.	218.	2
60	X29	TRAIN	0	1	199.	27.9	104.	180.	200.	218.	2
61	X30	TEST	0	1	175.	27.1	68.8	157.	176.	194.	2
62	X30	TRAIN	0	1	174.	27.6	75.9	155.	175.	195.	2
63	X31	TEST	0	1	177.	26.7	79.4	158.	178.	195.	2
64	X31	TRAIN	0	1	176.	26.5	78.2	158.	177.	195.	2
65	X32	TEST	0	1	179.	26.2	91.2	162.	180.	198.	2
66	X32	TRAIN	0	1	179.	26.7	82.7	161.	180.	197.	2
67	X33	TEST	0	1	176.	23.2	85.0	160.	177.	192.	2
68	X33	TRAIN	0	1	176.	23.5	79.2	159.	176.	192.	2
69	X34	TEST	0	1	172.	24.8	83.1	155.	172.	189.	2
70	X34	TRAIN	0	1	171.	25.0	87.9	154.	171.	188.	2
71	X35	TEST	0	1	191.	27.8	89.3	171.	192.	211.	2
72	X35	TRAIN	0	1	190.	28.5	95.4	170.	190.	211.	2
73	X36	TEST	0	1	172.	25.8	65.3	155.	173.	190.	2
74	X36	TRAIN	0	1	172.	25.6	84.1	154.	173.	190.	2
75	X37	TEST	0	1	170.	22.3	80.8	155.	170.	185.	2
76	X37	TRAIN	0	1	170.	22.3	85.2	155.	170.	185.	2
77	X38	TEST	0	1	182.	26.0	106.	164.	182.	200.	2
78	X38	TRAIN	0	1	181.	26.6	86.7	163.	182.	200.	2
79	X39	TEST	0	1	167.	24.0	74.9	151.	168.	184.	2
80	X39	TRAIN	0	1	167.	24.2	77.1	150.	167.	183.	2
81	X40	TEST	0	1	180.	26.8	84.4	162.	181.	198.	2
82	X40	TRAIN	0	1	179.	27.0	88.7	161.	180.	198.	2
83	X41	TEST	0	1	167.	24.6	65.8	150.	168.	184.	2
84	X41	TRAIN	0	1	167.	25.0	76.5	149.	167.	185.	2
85	X42	TEST	0	1	179.	26.6	69.9	161.	180.	197.	2
86	X42	TRAIN	0	1	178.	27.1	62.3	160.	179.	197.	2
87	X43	TEST	0	1	180.	26.7	91.8	163.	181.	199.	2
88	X43	TRAIN	0	1	179.	27.1	75.0	161.	181.	199.	2
89	X44	TEST	0	1	176.	26.6	76.2	159.	177.	195.	2
90	X44	TRAIN	0	1	175.	27.0	58.8	157.	176.	195.	2
91	X45	TEST	0	1	179.	24.3	87.8	164.	179.	196.	2
92	X45	TRAIN	0	1	178.	24.3	78.4	162.	179.	195.	2
93	X46	TEST	0	1	172.	25.1	82.6	155.	173.	189.	2
94	X46	TRAIN	0	1	171.	25.7	75.4	154.	172.	189.	2
95	X47	TEST	0	1	195.	26.7	90.8	178.	196.	214.	2
96	X47	TRAIN	0	1	195.	27.0	70.4	176.	196.	214.	2
97	X48	TEST	0	1	162.	25.5	72.3	144.	163.	179.	2
98	X48	TRAIN	0	1	161.	25.6	55.5	143.	161.	178.	2
99	X49	TEST	0	1	180.	25.0	87.7	163.	181.	197.	2
100	X49	TRAIN	0	1	179.	25.4	69.3	162.	180.	197.	2
101	X50	TEST	0	1	193.	26.2	101.	176.	194.	212.	2
102	X50	TRAIN	0	1	193.	26.2	82.4	175.	194.	211.	2
103	X51	TEST	0	1	175.	24.0	94.2	159.	175.	191.	2
104	X51	TRAIN	0	1	175.	24.1	82.6	158.	175.	192.	2
105	X52	TEST	0	1	177.	25.1	89.9	160.	177.	195.	2
106	X52	TRAIN	0	1	177.	25.4	84.9	160.	177.	194.	2
107	X53	TEST	0	1	188.	25.2	94.8	171.	189.	205.	2
108	X53	TRAIN	0	1	187.	25.2	89.1	170.	188.	205.	2

109	X54	TEST	0	1	164.	24.0	73.6	148.	165.	180.	2
110	X54	TRAIN	0	1	163.	24.0	63.3	147.	164.	180.	2
111	X55	TEST	0	1	175.	23.9	92.2	159.	175.	191.	2
112	X55	TRAIN	0	1	174.	24.1	85.5	158.	175.	191.	2
113	X56	TEST	0	1	191.	26.4	93.4	173.	191.	209.	2
114	X56	TRAIN	0	1	190.	26.6	90.6	172.	191.	208.	2
115	X57	TEST	0	1	176.	24.8	79.0	159.	176.	193.	2
116	X57	TRAIN	0	1	175.	25.3	78.7	158.	175.	192.	2
117	X58	TEST	0	1	185.	25.2	86.8	167.	186.	202.	2
118	X58	TRAIN	0	1	184.	25.7	98.0	166.	184.	202.	2
119	X59	TEST	0	1	168.	24.0	74.3	151.	168.	185.	2
120	X59	TRAIN	0	1	168.	23.9	85.0	151.	168.	184.	2
121	X60	TEST	0	1	175.	26.9	86.6	157.	176.	194.	2
122	X60	TRAIN	0	1	174.	27.5	69.5	155.	174.	194.	2
123	X61	TEST	0	1	178.	25.9	91.9	161.	178.	197.	2
124	X61	TRAIN	0	1	177.	26.2	70.1	160.	178.	196.	2
125	X62	TEST	0	1	177.	25.8	87.9	159.	177.	195.	2
126	X62	TRAIN	0	1	176.	26.2	84.6	158.	176.	194.	2
127	X63	TEST	0	1	163.	23.6	79.0	147.	164.	180.	2
128	X63	TRAIN	0	1	163.	23.7	81.7	147.	163.	179.	2
129	X64	TEST	0	1	177.	24.1	62.4	162.	178.	194.	2
130	X64	TRAIN	0	1	177.	24.5	70.2	160.	177.	194.	2
131	X65	TEST	0	1	176.	24.0	86.3	160.	177.	193.	2
132	X65	TRAIN	0	1	176.	24.0	89.4	160.	176.	192.	2
133	X66	TEST	0	1	177.	24.3	93.1	161.	178.	194.	2
134	X66	TRAIN	0	1	177.	24.8	91.9	160.	178.	195.	2
135	X67	TEST	0	1	179.	26.1	86.5	162.	180.	198.	2
136	X67	TRAIN	0	1	179.	26.6	80.5	160.	180.	198.	2
137	X68	TEST	0	1	160.	25.6	69.1	143.	161.	179.	2
138	X68	TRAIN	0	1	160.	26.0	69.6	141.	161.	178.	2
139	X69	TEST	0	1	159.	22.6	78.6	143.	159.	174.	2
140	X69	TRAIN	0	1	159.	22.9	64.9	143.	159.	175.	2
141	X70	TEST	0	1	167.	27.0	80.8	149.	169.	187.	2
142	X70	TRAIN	0	1	166.	27.4	61.1	148.	168.	186.	2
143	X71	TEST	0	1	179.	25.8	82.3	161.	179.	196.	2
144	X71	TRAIN	0	1	178.	26.2	97.8	160.	178.	196.	2
145	X72	TEST	0	1	165.	22.2	89.5	150.	165.	180.	2
146	X72	TRAIN	0	1	165.	22.2	80.7	150.	164.	180.	2
147	X73	TEST	0	1	177.	24.5	87.3	161.	178.	194.	2
148	X73	TRAIN	0	1	177.	24.6	84.9	160.	177.	194.	2
149	X74	TEST	0	1	166.	24.8	75.9	149.	168.	184.	2
150	X74	TRAIN	0	1	166.	24.8	68.4	149.	166.	183.	2
151	X75	TEST	0	1	183.	23.3	93.4	167.	183.	199.	2
152	X75	TRAIN	0	1	183.	23.5	69.4	167.	183.	199.	2
153	X76	TEST	0	1	181.	26.3	79.0	163.	181.	199.	2
154	X76	TRAIN	0	1	180.	26.6	84.2	161.	180.	198.	2
155	X77	TEST	0	1	176.	25.1	80.4	159.	177.	193.	2
156	X77	TRAIN	0	1	176.	25.1	88.2	159.	177.	194.	2
157	X78	TEST	0	1	166.	24.0	67.7	150.	166.	182.	2
158	X78	TRAIN	0	1	165.	24.2	82.4	149.	166.	182.	2
159	X79	TEST	0	1	182.	26.1	94.3	165.	183.	200.	2

160	X79	TRAIN	0	1	181.	26.1	82.1	163.	182.	199.	2
161	X80	TEST	0	1	181.	26.1	95.4	164.	182.	200.	2
162	X80	TRAIN	0	1	181.	26.8	83.7	163.	181.	200.	2
163	X81	TEST	0	1	172.	27.2	79.9	154.	173.	192.	2
164	X81	TRAIN	0	1	172.	27.5	71.8	152.	174.	191.	2
165	X82	TEST	0	1	184.	25.6	102.	167.	185.	202.	2
166	X82	TRAIN	0	1	183.	25.7	92.3	166.	184.	201.	2
167	X83	TEST	0	1	184.	29.9	92.5	164.	186.	205.	2
168	X83	TRAIN	0	1	183.	30.4	68.7	162.	185.	205.	2
169	X84	TEST	0	1	185.	26.9	89.0	166.	185.	204.	2
170	X84	TRAIN	0	1	184.	27.4	73.9	165.	184.	203.	2
171	X85	TEST	0	1	188.	27.5	91.1	170.	190.	208.	2
172	X85	TRAIN	0	1	188.	27.8	90.7	168.	189.	207.	2
173	X86	TEST	0	1	166.	22.9	85.2	151.	167.	182.	2
174	X86	TRAIN	0	1	165.	23.3	79.5	150.	166.	181.	2
175	X87	TEST	0	1	172.	26.7	81.0	154.	173.	191.	2
176	X87	TRAIN	0	1	171.	27.0	77.6	153.	173.	190.	2
177	X88	TEST	0	1	169.	24.0	81.3	153.	170.	186.	2
178	X88	TRAIN	0	1	168.	24.2	83.8	152.	169.	185.	2
179	X89	TEST	0	1	180.	24.9	90.6	163.	180.	197.	2
180	X89	TRAIN	0	1	179.	25.5	89.4	162.	179.	197.	2
181	X90	TEST	0	1	186.	26.2	103.	168.	186.	205.	2
182	X90	TRAIN	0	1	185.	26.5	79.8	167.	186.	204.	2
183	X91	TEST	0	1	167.	26.8	61.6	149.	169.	186.	2
184	X91	TRAIN	0	1	167.	27.5	69.4	147.	168.	186.	2
185	X92	TEST	0	1	172.	22.3	81.4	157.	172.	187.	2
186	X92	TRAIN	0	1	172.	22.1	88.9	157.	172.	187.	2
187	X93	TEST	0	1	194.	26.7	97.2	176.	195.	213.	2
188	X93	TRAIN	0	1	193.	26.6	97.8	176.	194.	212.	2
189	X94	TEST	0	1	177.	25.5	89.6	160.	177.	195.	2
190	X94	TRAIN	0	1	176.	26.1	89.4	159.	176.	194.	2
191	X95	TEST	0	1	182.	30.9	73.5	161.	184.	205.	2
192	X95	TRAIN	0	1	181.	31.4	73.0	159.	183.	204.	2
193	X96	TEST	0	1	173.	23.2	87.0	157.	173.	189.	2
194	X96	TRAIN	0	1	172.	23.1	88.7	157.	172.	188.	2
195	X97	TEST	0	1	196.	26.2	106.	178.	196.	215.	2
196	X97	TRAIN	0	1	196.	26.1	92.2	178.	196.	214.	2
197	X98	TEST	0	1	175.	27.1	77.7	155.	175.	194.	2
198	X98	TRAIN	0	1	174.	27.3	80.8	154.	174.	193.	2
199	X99	TEST	0	1	168.	23.4	93.1	152.	168.	184.	2
200	X99	TRAIN	0	1	167.	23.8	86.6	151.	168.	184.	2
201	X100	TEST	0	1	189.	28.3	72.8	170.	189.	208.	2
202	X100	TRAIN	0	1	188.	29.0	87.4	167.	188.	208.	2
203	X101	TEST	0	1	181.	25.1	88.5	165.	182.	198.	2
204	X101	TRAIN	0	1	181.	25.2	60.3	164.	182.	198.	2
205	X102	TEST	0	1	182.	26.7	82.0	163.	183.	201.	2
206	X102	TRAIN	0	1	181.	27.1	82.6	162.	182.	200.	2
207	X103	TEST	0	1	190.	26.5	81.9	172.	191.	209.	2
208	X103	TRAIN	0	1	189.	26.7	72.6	171.	190.	207.	2
209	X104	TEST	0	1	187.	26.5	91.0	170.	188.	206.	2
210	X104	TRAIN	0	1	187.	26.7	89.1	169.	187.	205.	2

211	X105	TEST	0	1	170.	27.2	69.0	151.	171.	189.	2
212	X105	TRAIN	0	1	169.	27.6	70.5	150.	170.	188.	2
213	X106	TEST	0	1	184.	24.6	95.8	168.	184.	202.	2
214	X106	TRAIN	0	1	184.	24.9	87.5	167.	184.	201.	2
215	X107	TEST	0	1	167.	24.3	74.0	150.	167.	185.	2
216	X107	TRAIN	0	1	166.	24.9	80.3	149.	167.	184.	2
217	X108	TEST	0	1	161.	24.0	69.7	144.	162.	178.	2
218	X108	TRAIN	0	1	160.	24.0	68.5	143.	161.	177.	2
219	X109	TEST	0	1	178.	25.3	86.8	161.	178.	195.	2
220	X109	TRAIN	0	1	177.	25.8	78.1	160.	178.	195.	2
221	X110	TEST	0	1	176.	25.1	86.7	160.	177.	194.	2
222	X110	TRAIN	0	1	175.	25.2	84.3	159.	176.	193.	2
223	X111	TEST	0	1	190.	24.9	105.	173.	190.	207.	2
224	X111	TRAIN	0	1	190.	24.7	99.5	173.	190.	206.	2
225	X112	TEST	0	1	171.	25.7	89.8	153.	172.	189.	2
226	X112	TRAIN	0	1	170.	25.7	77.1	153.	170.	187.	2
227	X113	TEST	0	1	164.	24.0	76.4	147.	164.	180.	2
228	X113	TRAIN	0	1	163.	24.3	76.4	146.	164.	180.	2
229	X114	TEST	0	1	177.	25.1	86.4	160.	176.	194.	2
230	X114	TRAIN	0	1	176.	25.4	65.6	159.	176.	194.	2
231	X115	TEST	0	1	197.	27.2	104.	179.	198.	216.	2
232	X115	TRAIN	0	1	196.	27.5	77.6	177.	198.	216.	2
233	X116	TEST	0	1	172.	26.2	72.6	155.	173.	191.	2
234	X116	TRAIN	0	1	171.	26.3	65.3	154.	173.	190.	2
235	X117	TEST	0	1	167.	23.6	83.7	150.	167.	183.	2
236	X117	TRAIN	0	1	166.	23.5	63.9	151.	167.	182.	2
237	X118	TEST	0	1	173.	24.0	85.4	157.	173.	190.	2
238	X118	TRAIN	0	1	172.	24.3	71.6	155.	173.	189.	2
239	X119	TEST	0	1	171.	22.7	90.9	155.	171.	186.	2
240	X119	TRAIN	0	1	170.	23.0	81.2	154.	171.	186.	2
241	X120	TEST	0	1	185.	28.8	87.7	165.	186.	205.	2
242	X120	TRAIN	0	1	184.	29.0	85.2	164.	185.	204.	2
243	X121	TEST	0	1	170.	25.1	77.5	153.	170.	188.	2
244	X121	TRAIN	0	1	170.	25.5	78.1	152.	170.	187.	2
245	X122	TEST	0	1	184.	24.2	93.8	168.	184.	200.	2
246	X122	TRAIN	0	1	183.	24.1	95.9	166.	184.	200.	2
247	X123	TEST	0	1	181.	25.8	86.6	163.	181.	199.	2
248	X123	TRAIN	0	1	180.	26.4	78.0	162.	181.	198.	2
249	X124	TEST	0	1	176.	23.5	80.9	161.	177.	192.	2
250	X124	TRAIN	0	1	176.	23.6	74.7	160.	176.	192.	2
251	X125	TEST	0	1	186.	27.0	96.4	168.	187.	205.	2
252	X125	TRAIN	0	1	186.	27.1	78.9	167.	187.	204.	2
253	X126	TEST	0	1	185.	23.8	96.1	169.	185.	202.	2
254	X126	TRAIN	0	1	184.	24.1	92.9	168.	184.	201.	2
255	X127	TEST	0	1	176.	25.0	91.1	158.	176.	193.	2
256	X127	TRAIN	0	1	175.	25.3	74.0	158.	175.	192.	2
257	X128	TEST	0	1	168.	25.2	73.7	150.	168.	186.	2
258	X128	TRAIN	0	1	167.	25.6	67.0	149.	167.	185.	2
259	X129	TEST	0	1	172.	25.7	82.4	155.	173.	190.	2
260	X129	TRAIN	0	1	171.	25.7	75.2	154.	172.	190.	2
261	X130	TEST	0	1	181.	25.3	95.5	164.	182.	199.	2

262	X130	TRAIN	0	1	181.	25.3	87.3	164.	181.	198.	2
263	X131	TEST	0	1	178.	22.8	98.8	163.	178.	193.	2
264	X131	TRAIN	0	1	177.	22.7	87.4	162.	178.	193.	2
265	X132	TEST	0	1	167.	24.6	82.4	150.	168.	184.	2
266	X132	TRAIN	0	1	166.	25.0	46.8	149.	167.	183.	2
267	X133	TEST	0	1	191.	27.1	102.	172.	192.	210.	2
268	X133	TRAIN	0	1	190.	27.8	86.6	170.	191.	209.	2
269	X134	TEST	0	1	179.	25.7	85.4	162.	180.	197.	2
270	X134	TRAIN	0	1	179.	25.8	92.7	161.	180.	197.	2
271	X135	TEST	0	1	186.	26.0	94.3	169.	186.	204.	2
272	X135	TRAIN	0	1	185.	26.4	95.4	167.	185.	204.	2
273	X136	TEST	0	1	190.	26.8	79.0	172.	191.	209.	2
274	X136	TRAIN	0	1	190.	27.2	89.5	171.	191.	209.	2
275	X137	TEST	0	1	164.	23.0	77.3	148.	164.	180.	2
276	X137	TRAIN	0	1	163.	23.3	54.6	148.	163.	179.	2
277	X138	TEST	0	1	194.	27.9	95.1	174.	194.	213.	2
278	X138	TRAIN	0	1	193.	28.4	91.2	174.	195.	213.	2
279	X139	TEST	0	1	169.	24.1	81.8	153.	169.	186.	2
280	X139	TRAIN	0	1	169.	24.3	78.6	152.	169.	186.	2
281	X140	TEST	0	1	182.	26.3	87.5	165.	183.	201.	2
282	X140	TRAIN	0	1	182.	26.7	72.5	163.	183.	200.	2

Measuring Association

```
set.seed(1234)
corr.mat <- round(cor(Data.df %>% select(-c(Y, Subset))), 3)
corr.thresh <- which(abs(corr.mat) > 0.85 & abs(corr.mat) < 1, arr.ind = TRUE)
print(corr.thresh)
```

	row	col
X32	32	2
X120	120	2
X7	7	3
X31	31	3
X40	40	3
X129	129	3
X100	100	5
X103	103	5
X104	104	5
X14	14	6
X19	19	6
X3	3	7
X18	18	7
X24	24	7
X82	82	7
X91	91	7
X103	103	7
X120	120	7
X14	14	9
X15	15	9

X18	18	9
X29	29	9
X35	35	9
X87	87	9
X95	95	9
X103	103	9
X104	104	9
X120	120	9
X98	98	11
X100	100	11
X103	103	11
X45	45	12
X47	47	12
X49	49	12
X53	53	12
X66	66	12
X88	88	12
X101	101	12
X103	103	12
X120	120	12
X122	122	12
X38	38	13
X119	119	13
X6	6	14
X9	9	14
X15	15	14
X18	18	14
X19	19	14
X29	29	14
X30	30	14
X31	31	14
X35	35	14
X42	42	14
X43	43	14
X44	44	14
X47	47	14
X50	50	14
X61	61	14
X68	68	14
X70	70	14
X81	81	14
X83	83	14
X85	85	14
X87	87	14
X90	90	14
X91	91	14
X95	95	14
X102	102	14
X103	103	14
X104	104	14
X105	105	14

X107	107	14
X109	109	14
X110	110	14
X112	112	14
X115	115	14
X120	120	14
X125	125	14
X129	129	14
X133	133	14
X134	134	14
X140	140	14
X9	9	15
X14	14	15
X18	18	15
X29	29	15
X50	50	15
X58	58	15
X79	79	15
X83	83	15
X105	105	15
X18	18	16
X113	113	16
X71	71	17
X100	100	17
X135	135	17
X7	7	18
X9	9	18
X14	14	18
X15	15	18
X16	16	18
X19	19	18
X24	24	18
X25	25	18
X29	29	18
X30	30	18
X31	31	18
X35	35	18
X42	42	18
X43	43	18
X44	44	18
X45	45	18
X47	47	18
X52	52	18
X54	54	18
X58	58	18
X66	66	18
X70	70	18
X74	74	18
X79	79	18
X82	82	18
X83	83	18

X91	91	18
X95	95	18
X101	101	18
X103	103	18
X104	104	18
X110	110	18
X118	118	18
X120	120	18
X125	125	18
X129	129	18
X133	133	18
X136	136	18
X138	138	18
X140	140	18
X6	6	19
X14	14	19
X18	18	19
X35	35	19
X42	42	19
X43	43	19
X44	44	19
X70	70	19
X81	81	19
X103	103	19
X120	120	19
X29	29	22
X55	55	23
X7	7	24
X18	18	24
X18	18	25
X47	47	25
X50	50	26
X82	82	26
X103	103	26
X9	9	29
X14	14	29
X15	15	29
X18	18	29
X22	22	29
X40	40	29
X47	47	29
X58	58	29
X61	61	29
X76	76	29
X82	82	29
X83	83	29
X85	85	29
X109	109	29
X110	110	29
X115	115	29
X125	125	29

X136	136	29
X14	14	30
X18	18	30
X54	54	30
X94	94	30
X95	95	30
X98	98	30
X100	100	30
X107	107	30
X3	3	31
X14	14	31
X18	18	31
X44	44	31
X47	47	31
X66	66	31
X79	79	31
X82	82	31
X90	90	31
X93	93	31
X97	97	31
X101	101	31
X115	115	31
X120	120	31
X138	138	31
X2	2	32
X79	79	32
X103	103	32
X115	115	32
X123	123	32
X95	95	34
X103	103	34
X120	120	34
X9	9	35
X14	14	35
X18	18	35
X19	19	35
X43	43	35
X50	50	35
X58	58	35
X83	83	35
X85	85	35
X90	90	35
X93	93	35
X102	102	35
X120	120	35
X129	129	35
X133	133	35
X134	134	35
X138	138	35
X47	47	36
X74	74	36

X103	103	36
X120	120	36
X13	13	38
X43	43	38
X3	3	40
X29	29	40
X70	70	40
X77	77	40
X82	82	40
X85	85	40
X112	112	40
X120	120	40
X138	138	40
X81	81	41
X95	95	41
X14	14	42
X18	18	42
X19	19	42
X81	81	42
X14	14	43
X18	18	43
X19	19	43
X35	35	43
X38	38	43
X81	81	43
X85	85	43
X91	91	43
X112	112	43
X129	129	43
X138	138	43
X14	14	44
X18	18	44
X19	19	44
X31	31	44
X45	45	44
X47	47	44
X50	50	44
X70	70	44
X79	79	44
X83	83	44
X103	103	44
X115	115	44
X129	129	44
X138	138	44
X12	12	45
X18	18	45
X44	44	45
X47	47	45
X115	115	45
X50	50	46
X90	90	46

X101	101	46
X103	103	46
X12	12	47
X14	14	47
X18	18	47
X25	25	47
X29	29	47
X31	31	47
X36	36	47
X44	44	47
X45	45	47
X70	70	47
X74	74	47
X75	75	47
X101	101	47
X115	115	47
X129	129	47
X136	136	47
X138	138	47
X12	12	49
X14	14	50
X15	15	50
X26	26	50
X35	35	50
X44	44	50
X46	46	50
X115	115	50
X136	136	50
X138	138	50
X18	18	52
X82	82	52
X103	103	52
X12	12	53
X103	103	53
X18	18	54
X30	30	54
X103	103	54
X23	23	55
X15	15	58
X18	18	58
X29	29	58
X35	35	58
X82	82	58
X129	129	58
X61	61	60
X100	100	60
X102	102	60
X14	14	61
X29	29	61
X60	60	61
X95	95	62

X98	98	62
X103	103	62
X107	107	62
X66	66	63
X75	75	64
X12	12	66
X18	18	66
X31	31	66
X63	63	66
X90	90	66
X93	93	66
X107	107	66
X120	120	66
X138	138	66
X70	70	67
X138	138	67
X14	14	68
X95	95	68
X116	116	68
X133	133	68
X14	14	70
X18	18	70
X19	19	70
X40	40	70
X44	44	70
X47	47	70
X67	67	70
X81	81	70
X83	83	70
X110	110	70
X115	115	70
X134	134	70
X17	17	71
X100	100	71
X82	82	73
X93	93	73
X110	110	73
X18	18	74
X36	36	74
X47	47	74
X83	83	74
X90	90	74
X95	95	74
X103	103	74
X120	120	74
X47	47	75
X64	64	75
X29	29	76
X79	79	76
X104	104	76
X40	40	77

X138	138	77
X15	15	79
X18	18	79
X31	31	79
X32	32	79
X44	44	79
X76	76	79
X81	81	79
X93	93	79
X115	115	79
X123	123	79
X100	100	80
X103	103	80
X14	14	81
X19	19	81
X41	41	81
X42	42	81
X43	43	81
X70	70	81
X79	79	81
X91	91	81
X127	127	81
X138	138	81
X7	7	82
X18	18	82
X26	26	82
X29	29	82
X31	31	82
X40	40	82
X52	52	82
X58	58	82
X73	73	82
X90	90	82
X97	97	82
X111	111	82
X115	115	82
X125	125	82
X127	127	82
X129	129	82
X136	136	82
X14	14	83
X15	15	83
X18	18	83
X29	29	83
X35	35	83
X44	44	83
X70	70	83
X74	74	83
X95	95	83
X105	105	83
X118	118	83

X133	133	83
X14	14	85
X29	29	85
X35	35	85
X40	40	85
X43	43	85
X87	87	85
X112	112	85
X138	138	85
X9	9	87
X14	14	87
X85	85	87
X95	95	87
X120	120	87
X12	12	88
X103	103	88
X14	14	90
X31	31	90
X35	35	90
X46	46	90
X66	66	90
X74	74	90
X82	82	90
X91	91	90
X101	101	90
X103	103	90
X105	105	90
X120	120	90
X125	125	90
X138	138	90
X7	7	91
X14	14	91
X18	18	91
X43	43	91
X81	81	91
X90	90	91
X120	120	91
X140	140	91
X31	31	93
X35	35	93
X66	66	93
X73	73	93
X79	79	93
X101	101	93
X103	103	93
X30	30	94
X100	100	94
X9	9	95
X14	14	95
X18	18	95
X30	30	95

```

X34    34  95
X41    41  95
X62    62  95
X68    68  95
X74    74  95
X83    83  95
X87    87  95
X103  103  95
X105  105  95
X120  120  95
X125  125  95
X133  133  95
X31    31  97
X82    82  97
X11    11  98
X30    30  98
X62    62  98
X100  100  98
X103  103  98
X123  123  98
X128  128  98
[ reached getOption("max.print") -- omitted 216 rows ]
print(corr.mat[corr.thresh])
[1] 0.851 0.867 0.853 0.852 0.852 0.853 0.854 0.872 0.863 0.865 0.852 0.853 0.865 0.868 0
[21] 0.869 0.860 0.865 0.853 0.853 0.864 0.855 0.867 0.854 0.860 0.858 0.855 0.852 0.860 0
[41] 0.857 0.854 0.854 0.865 0.859 0.860 0.880 0.890 0.878 0.869 0.852 0.863 0.871 0.863 0
[61] 0.870 0.879 0.872 0.868 0.865 0.868 0.854 0.888 0.851 0.856 0.854 0.859 0.857 0.862 0
[81] 0.852 0.851 0.883 0.860 0.856 0.852 0.859 0.856 0.874 0.854 0.875 0.858 0.852 0.858 0
[101] 0.858 0.858 0.861 0.856 0.864 0.853 0.858 0.855 0.858 0.857 0.873 0.861 0.889 0.857 0
[121] 0.869 0.857 0.859 0.852 0.860 0.866 0.855 0.851 0.854 0.864 0.854 0.854 0.857 0.859 0
[141] 0.865 0.851 0.851 0.872 0.860 0.852 0.853 0.851 0.853 0.868 0.861 0.856 0.851 0.853 0
[161] 0.851 0.861 0.851 0.854 0.862 0.853 0.860 0.853 0.871 0.851 0.856 0.858 0.881 0.860 0
[181] 0.893 0.859 0.852 0.852 0.858 0.877 0.858 0.858 0.875 0.865 0.863 0.860 0.854 0.851 0
[201] 0.861 0.858 0.856 0.861 0.853 0.865 0.863 0.855 0.856 0.864 0.855 0.864 0.857 0.857 0
[221] 0.852 0.869 0.860 0.851 0.860 0.857 0.854 0.852 0.852 0.861 0.852 0.857 0.856 0.862 0
[241] 0.858 0.865 0.856 0.863 0.857 0.851 0.864 0.852 0.860 0.852 0.858 0.854 0.865 0.862 0
[261] 0.862 0.851 0.853 0.852 0.851 0.872 0.871 0.866 0.855 0.861 0.856 0.863 0.853 0.852 0
[281] 0.851 0.851 0.858 0.860 0.866 0.863 0.860 0.856 0.855 0.858 0.865 0.852 0.856 0.875 0
[301] 0.852 0.855 0.852 0.869 0.857 0.851 0.854 0.852 0.853 0.855 0.857 0.855 0.853 0.856 0
[321] 0.864 0.855 0.880 0.862 0.855 0.857 0.861 0.866 0.852 0.858 0.857 0.863 0.861 0.858 0
[341] 0.853 0.859 0.861 0.853 0.859 0.858 0.889 0.856 0.872 0.852 0.851 0.860 0.853 0.853 0
[361] 0.854 0.854 0.853 0.854 0.851 0.856 0.857 0.860 0.860 0.864 0.860 0.855 0.857 0.853 0
[381] 0.875 0.863 0.853 0.857 0.852 0.851 0.859 0.855 0.855 0.857 0.870 0.860 0.860 0.856 0
[401] 0.865 0.869 0.852 0.860 0.865 0.856 0.851 0.858 0.854 0.855 0.854 0.862 0.866 0.872 0
[421] 0.853 0.857 0.852 0.871 0.857 0.868 0.853 0.856 0.864 0.872 0.871 0.857 0.862 0.852 0
[441] 0.865 0.874 0.863 0.856 0.865 0.863 0.853 0.853 0.856 0.860 0.855 0.854 0.863 0.853 0
[461] 0.859 0.858 0.865 0.854 0.865 0.867 0.860 0.870 0.867 0.854 0.851 0.865 0.851 0.862 0
[481] 0.854 0.857 0.853 0.860 0.868 0.865 0.860 0.863 0.874 0.865 0.867 0.854 0.854 0.854 0
[501] 0.854 0.860 0.858 0.893 0.864 0.866 0.855 0.860 0.886 0.854 0.852 0.856 0.865 0.855 0
[521] 0.865 0.852 0.855 0.852 0.888 0.856 0.855 0.857 0.854 0.872 0.863 0.864 0.858 0.858 0
[541] 0.860 0.851 0.853 0.854 0.853 0.855 0.866 0.864 0.857 0.856 0.853 0.851 0.860 0.868 0

```



```
[561] 0.855 0.870 0.863 0.852 0.863 0.855 0.856 0.855 0.853 0.852 0.855 0.868 0.852 0.858 0
[581] 0.852 0.859 0.859 0.859 0.852 0.870 0.857 0.858 0.857 0.851 0.862 0.851 0.856 0.851 0
[601] 0.867 0.852 0.863 0.867 0.858 0.851 0.861 0.872 0.853 0.865 0.855 0.858 0.859 0.866 0
[621] 0.856 0.855 0.854 0.867 0.852 0.867 0.853 0.860 0.864 0.853 0.851 0.853 0.857 0.857 0
[641] 0.874 0.870 0.858 0.859 0.866 0.857 0.858 0.855 0.854 0.856 0.854 0.881 0.872 0.854 0
[661] 0.857 0.865 0.853 0.856 0.854 0.867 0.865 0.871 0.852 0.853 0.854 0.856 0.865 0.857 0
[681] 0.852 0.852 0.855 0.856 0.851 0.870 0.859 0.860 0.856 0.852 0.861 0.856 0.851 0.853 0
[701] 0.867 0.859 0.858 0.855 0.863 0.852 0.852 0.873 0.866 0.851 0.865 0.851 0.858 0.867 0
```

```
digits <- 3
significance.level <- 0.85

# Get numeric data.
Data.numeric <- Data.df %>% select(-c(Y, Subset))

# Calculate correlations and drop ones below threshold.
corr.Data <- round(cor(Data.numeric), digits)

# Drop duplicates and correlations of 1.
corr.Data[lower.tri(corr.Data, diag = TRUE)] <- NA
corr.Data[corr.Data == 1] <- NA

# Turn into 3-column table.
corr.Data %<>% as.table() %>% as.data.frame()

# Select significant values.
corr.Data <- subset(corr.Data, abs(Freq) > significance.level)
corr.Data <- corr.Data[order(-abs(corr.Data$Freq)),]

# Omit NA values.
corr.Data %<>% na.omit()
print(corr.Data)
  Var1 Var2 Freq
13890 X30 X100 0.893
2534 X14 X19 0.890
6458 X18 X47 0.889
9674 X14 X70 0.889
14154 X14 X102 0.888
13958 X98 X100 0.886
1969 X9 X15 0.883
17389 X29 X125 0.881
2394 X14 X18 0.880
8414 X14 X61 0.880
11494 X14 X83 0.879
3934 X14 X29 0.878
6051 X31 X44 0.877
10951 X31 X79 0.875
14575 X15 X105 0.875
19227 X47 X138 0.875
10935 X15 X79 0.874
16747 X87 X120 0.874
```

16755	X95	X120	0.874
6038	X18	X44	0.873
19295	X115	X138	0.873
9679	X19	X70	0.872
11774	X14	X85	0.872
14285	X5	X103	0.872
16004	X44	X115	0.872
17442	X82	X125	0.872
5754	X14	X42	0.871
11550	X70	X83	0.871
11789	X29	X85	0.871
17964	X44	X129	0.871
11214	X14	X81	0.870
12915	X35	X93	0.870
14906	X66	X107	0.870
16763	X103	X120	0.870
18777	X17	X135	0.870
2389	X9	X18	0.869
4074	X14	X30	0.869
11358	X18	X82	0.869
19215	X35	X138	0.869
19230	X50	X138	0.869
3227	X7	X24	0.868
12054	X14	X87	0.868
12614	X14	X91	0.868
13243	X83	X95	0.868
14378	X98	X103	0.868
14523	X103	X104	0.868
7998	X18	X58	0.867
12946	X66	X93	0.867
15625	X85	X112	0.867
15974	X14	X115	0.867
16662	X2	X120	0.867
16669	X9	X120	0.867
17955	X35	X129	0.867
18575	X95	X133	0.867
19246	X66	X138	0.867
19551	X91	X140	0.867
6484	X44	X47	0.866
13931	X71	X100	0.866
14298	X18	X103	0.866
14342	X62	X103	0.866
16042	X82	X115	0.866
19224	X44	X138	0.866
19300	X120	X138	0.866
1826	X6	X14	0.865
2387	X7	X18	0.865
4769	X9	X35	0.865
5759	X19	X42	0.865
11347	X7	X82	0.865
11371	X31	X82	0.865

12474	X14	X90	0.865
12681	X81	X91	0.865
13247	X87	X95	0.865
14093	X93	X101	0.865
16007	X47	X115	0.865
16751	X91	X120	0.865
17455	X95	X125	0.865
17880	X100	X128	0.865
17963	X43	X129	0.865
18494	X14	X133	0.865
19598	X138	X140	0.865
3938	X18	X29	0.864
5915	X35	X43	0.864
8015	X35	X58	0.864
12607	X7	X91	0.864
13920	X60	X100	0.864
14289	X9	X103	0.864
14354	X74	X103	0.864
16678	X18	X120	0.864
16750	X90	X120	0.864
18563	X83	X133	0.864
4774	X14	X35	0.863
5894	X14	X43	0.863
6485	X45	X47	0.863
6874	X14	X50	0.863
9112	X12	X66	0.863
10952	X32	X79	0.863
12192	X12	X88	0.863
12491	X31	X90	0.863
14012	X12	X101	0.863
14090	X90	X101	0.863
14287	X7	X103	0.863
14425	X5	X104	0.863
14655	X95	X105	0.863
15554	X14	X112	0.863
15923	X103	X114	0.863
18583	X103	X133	0.863
19265	X85	X138	0.863
6904	X44	X50	0.862
8429	X29	X61	0.862
10938	X18	X79	0.862
11800	X40	X85	0.862
13050	X30	X94	0.862
15274	X14	X110	0.862
15482	X82	X111	0.862
19220	X40	X138	0.862
19223	X43	X138	0.862
3238	X18	X24	0.861
5489	X29	X40	0.861
6178	X18	X45	0.861
9118	X18	X66	0.861

9394	X14	X68	0.861
13642	X62	X98	0.861
14314	X34	X103	0.861
15992	X32	X115	0.861
18982	X82	X136	0.861
1974	X14	X15	0.860
3929	X9	X29	0.860
6476	X36	X47	0.860
6732	X12	X49	0.860
9707	X47	X70	0.860
11219	X19	X81	0.860
11241	X41	X81	0.860
11243	X43	X81	0.860
11369	X29	X82	0.860
12534	X74	X90	0.860
12911	X31	X93	0.860
13234	X74	X95	0.860
13610	X30	X98	0.860
13871	X11	X100	0.860
13954	X94	X100	0.860
14018	X18	X101	0.860
14316	X36	X103	0.860
14375	X95	X103	0.860
16674	X14	X120	0.860
16734	X74	X120	0.860
17464	X104	X125	0.860
18929	X29	X136	0.860
1829	X9	X14	0.859
6875	X15	X50	0.859
12618	X18	X91	0.859
14854	X14	X107	0.859
14870	X30	X107	0.859
16039	X79	X115	0.859
16168	X68	X116	0.859
16765	X105	X120	0.859
17722	X82	X127	0.859
18918	X18	X136	0.859
19247	X67	X138	0.859
2396	X16	X18	0.858
2538	X18	X19	0.858
4218	X18	X31	0.858
5758	X18	X42	0.858
6471	X31	X47	0.858
9131	X31	X66	0.858
9163	X63	X66	0.858
9817	X17	X71	0.858
11398	X58	X82	0.858
12643	X43	X91	0.858
13877	X17	X100	0.858
14047	X47	X101	0.858
14291	X11	X103	0.858

14292	X12	X103	0.858
14943	X103	X107	0.858
15989	X29	X115	0.858
16030	X70	X115	0.858
16764	X104	X120	0.858
17112	X32	X123	0.858
17721	X81	X127	0.858
18548	X68	X133	0.858
19257	X77	X138	0.858
19478	X18	X140	0.858
5898	X18	X43	0.857
6454	X14	X47	0.857
7158	X18	X52	0.857
7450	X30	X54	0.857
10424	X64	X75	0.857
10680	X40	X77	0.857
10996	X76	X79	0.857
11498	X18	X83	0.857
11515	X35	X83	0.857
11554	X74	X83	0.857
11795	X35	X85	0.857
13222	X62	X95	0.857
14360	X80	X103	0.857
14942	X102	X107	0.857
15134	X14	X109	0.857
16695	X35	X120	0.857
16696	X36	X120	0.857
16952	X12	X122	0.857
17878	X98	X128	0.857
18498	X18	X133	0.857
2395	X15	X18	0.856
3378	X18	X25	0.856
4779	X19	X35	0.856
6034	X14	X44	0.856
6204	X44	X45	0.856
7995	X15	X58	0.856
9678	X18	X70	0.856
10267	X47	X74	0.856
11242	X42	X81	0.856
11380	X40	X82	0.856
12526	X66	X90	0.856
13194	X34	X95	0.856
14175	X35	X102	0.856
14368	X88	X103	0.856
14434	X14	X104	0.856
15289	X29	X110	0.856
16463	X83	X118	0.856
16700	X40	X120	0.856
17374	X14	X125	0.856
17740	X100	X127	0.856
17934	X14	X129	0.856

18749	X129	X134	0.856
18947	X47	X136	0.856
19596	X136	X140	0.856
4778	X18	X35	0.855
6172	X12	X45	0.855
6895	X35	X50	0.855
7438	X18	X54	0.855
8460	X60	X61	0.855
10407	X47	X75	0.855
12542	X82	X90	0.855
13940	X80	X100	0.855
14200	X60	X102	0.855
14334	X54	X103	0.855
14429	X9	X104	0.855
14438	X18	X104	0.855
14521	X101	X104	0.855
16010	X50	X115	0.855
16483	X103	X118	0.855
17159	X79	X123	0.855
18580	X100	X133	0.855
18690	X70	X134	0.855
19261	X81	X138	0.855
5193	X13	X38	0.854
8009	X29	X58	0.854
10238	X18	X74	0.854
11413	X73	X82	0.854
11495	X15	X83	0.854
12690	X90	X91	0.854
12953	X73	X93	0.854
13174	X14	X95	0.854
13201	X41	X95	0.854
13471	X31	X97	0.854
13522	X82	X97	0.854
13591	X11	X98	0.854
13865	X5	X100	0.854
14312	X32	X103	0.854
14332	X52	X103	0.854
14380	X100	X103	0.854
14574	X14	X105	0.854
15583	X43	X112	0.854
16398	X18	X118	0.854
16533	X13	X119	0.854
17178	X98	X123	0.854
17378	X18	X125	0.854
17450	X90	X125	0.854
17462	X102	X125	0.854
17938	X18	X129	0.854
18002	X82	X129	0.854
18515	X35	X133	0.854
843	X3	X7	0.853
4078	X18	X30	0.853

6886	X26	X50	0.853
7583	X23	X55	0.853
9727	X67	X70	0.853
10529	X29	X76	0.853
10964	X44	X79	0.853
11270	X70	X81	0.853
11509	X29	X83	0.853
12049	X9	X87	0.853
12495	X35	X90	0.853
12506	X46	X90	0.853
13169	X9	X95	0.853
13228	X68	X95	0.853
14046	X46	X101	0.853
14326	X46	X103	0.853
14333	X53	X103	0.853
14370	X90	X103	0.853
14496	X76	X104	0.853
14643	X83	X105	0.853
15333	X73	X110	0.853
15580	X40	X112	0.853
16005	X45	X115	0.853
16672	X12	X120	0.853
16679	X19	X120	0.853
16694	X34	X120	0.853
17923	X3	X129	0.853
17978	X58	X129	0.853
19211	X31	X138	0.853
2526	X6	X19	0.852
3935	X15	X29	0.852
4203	X3	X31	0.852
4214	X14	X31	0.852
5463	X3	X40	0.852
5918	X38	X43	0.852
6452	X12	X47	0.852
6906	X46	X50	0.852
7292	X12	X53	0.852
9700	X40	X70	0.852
11279	X79	X81	0.852
11366	X26	X82	0.852
11524	X44	X83	0.852
11803	X43	X85	0.852
13178	X18	X95	0.852
13190	X30	X95	0.852
14299	X19	X103	0.852
14381	X101	X103	0.852
14520	X100	X104	0.852
14650	X90	X105	0.852
14664	X104	X105	0.852
14902	X62	X107	0.852
15696	X16	X113	0.852
16061	X101	X115	0.852

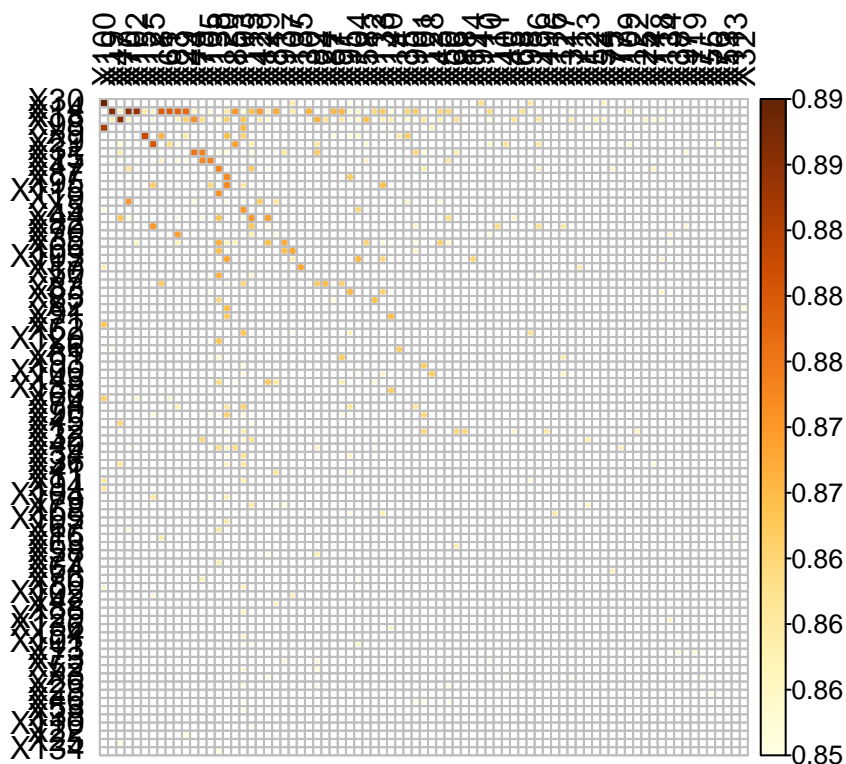

```

16063 X103 X115 0.852
16070 X110 X115 0.852
16667 X7 X120 0.852
16726 X66 X120 0.852
17967 X47 X129 0.852
18634 X14 X134 0.852
18655 X35 X134 0.852
18950 X50 X136 0.852
19270 X90 X138 0.852
[ reached 'max' / getOption("max.print") -- omitted 25 rows ]
# Convert back to matrix.
corr.Mat <- tidyr::pivot_wider(corr.Data, id_cols = Var1, names_from = Var2, values_from = 1)
  data.frame(row.names = 1) %>%
  as.matrix()
print(corr.Mat)
      X100  X19  X47  X70  X102  X15  X125  X18  X61  X83  X29  X44  X79  X105  X
X30  0.893   NA   NA   NA   NA   NA   NA   NA   NA   NA   NA   NA   NA   NA
X14   NA  0.890  0.857  0.889  0.888  0.860  0.856  0.880  0.880  0.879  0.878  0.856   NA  0.854
X18   NA  0.858  0.889  0.856   NA   NA  0.854   NA   NA  0.857  0.864  0.873  0.862   NA  0.8
X98  0.886   NA   NA   NA   NA   NA   NA   NA   NA   NA   NA   NA   NA   NA
X9    NA   NA   NA   NA   NA   NA  0.883   NA  0.869   NA   NA  0.860   NA   NA
X29   NA   NA  0.851   NA   NA   NA  0.881   NA  0.862  0.853   NA   NA   NA   NA
X31   NA   NA  0.858   NA   NA   NA   NA   NA   NA   NA   NA   NA  0.877  0.875   NA  0.8
X15   NA   NA   NA   NA   NA   NA   NA   NA  0.856   NA  0.854  0.852   NA  0.874  0.875
X47   NA   NA   NA  0.860   NA   NA   NA   NA   NA   NA   NA   NA   NA   NA   NA  0.8
X87   NA   NA   NA   NA   NA   NA   NA   NA   NA   NA   NA   NA   NA   NA   NA
X95   NA   NA   NA   NA   NA   NA   NA  0.865   NA   NA   NA   NA   NA   NA  0.863
X115  NA   NA   NA   NA   NA   NA   NA   NA   NA   NA   NA   NA   NA   NA   NA  0.8
      X129  X81  X93  X107  X135  X30  X82  X24  X87  X91  X95  X104  X58  X112  X13
X30   NA   NA   NA  0.859   NA   NA   NA   NA   NA   NA   NA  0.852   NA   NA   NA
X14  0.856  0.870   NA  0.859   NA  0.869   NA   NA  0.868  0.868  0.854  0.856   NA  0.863  0.86
X18  0.854   NA   NA   NA   NA  0.853  0.869  0.861   NA  0.859  0.852  0.855  0.867   NA  0.86
X98   NA   NA   NA   NA   NA   NA   NA   NA   NA   NA   NA   NA   NA   NA   NA
X9    NA   NA   NA   NA   NA   NA   NA   NA   NA  0.853   NA  0.853  0.855   NA   NA
X29   NA   NA   NA   NA   NA   NA   NA  0.860   NA   NA   NA   NA   NA  0.854   NA
X31   NA   NA  0.860   NA   NA   NA  0.865   NA   NA   NA   NA   NA   NA   NA   NA
X15   NA   NA   NA   NA   NA   NA   NA   NA   NA   NA   NA   NA   NA  0.856   NA
X47  0.852   NA   NA   NA   NA   NA   NA   NA   NA   NA   NA   NA   NA   NA   NA
X87   NA   NA   NA   NA   NA   NA   NA   NA   NA   NA   NA  0.865   NA   NA   NA
X95   NA   NA   NA   NA   NA   NA   NA   NA   NA   NA   NA   NA   NA   NA   NA  0.86
X115  NA   NA   NA   NA   NA   NA   NA   NA   NA   NA   NA   NA   NA   NA   NA
      X128  X43  X50  X66  X88  X114  X94  X110  X111  X40  X45  X68  X98  X136  X4
X30   NA   NA   NA   NA   NA   NA  0.862   NA   NA   NA   NA   NA  0.860   NA   NA
X14   NA  0.863  0.863   NA   NA   NA   NA  0.862   NA   NA   NA  0.861   NA   NA
X18   NA  0.857   NA  0.861   NA   NA   NA  0.851   NA   NA  0.861   NA   NA  0.859
X98  0.857   NA   NA   NA   NA   NA   NA   NA   NA   NA   NA   NA   NA   NA
X9    NA   NA   NA   NA   NA   NA   NA   NA   NA   NA   NA   NA   NA   NA   NA
X29   NA   NA   NA   NA   NA   NA   NA  0.856   NA  0.861   NA   NA   NA  0.860
X31   NA   NA   NA  0.858   NA   NA   NA   NA   NA   NA   NA   NA   NA   NA   NA
X15   NA   NA  0.859   NA   NA   NA   NA   NA   NA   NA   NA   NA   NA   NA
X47   NA   NA   NA   NA   NA   NA   NA   NA   NA   NA   NA   NA   NA  0.856

```

X87	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
X95	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
X115	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
	X52	X54	X75	X77	X109	X122	X25	X74	X118	X134	X38	X97	X119	X7	X
X30	NA	0.857	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
X14	NA	NA	NA	NA	0.857	NA	NA	NA	NA	0.852	NA	NA	NA	NA	NA
X18	0.857	0.855	NA	NA	NA	NA	0.856	0.854	0.854	NA	NA	NA	NA	NA	NA
X98	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
X9	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
X29	NA	NA	NA	NA	0.851	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
X31	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	0.854	NA	NA	NA
X15	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
X47	NA	NA	0.855	NA	NA	NA	NA	0.856	NA	NA	NA	NA	NA	NA	NA
X87	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
X95	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
X115	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA

```
[ reached getOption("max.print") -- omitted 68 rows ]
corrplot::corrplot(corr.Mat, is.corr = FALSE, tl.col = "black", na.label = " ")
```



Model Baseline

```
# Fit linear model.
lm.1 <- lm(Y ~ ., Data.train)

# Print summary.
summary(lm.1) %>% print()
```

Call:

```
lm(formula = Y ~ ., data = Data.train)
```

Residuals:

Min	1Q	Median	3Q	Max
-2425.8	-297.7	-38.2	247.7	5937.1

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)	
(Intercept)	-6279.0045	82.3127	-76.282	< 2e-16	***
X1	-9.5274	15.9168	-0.599	0.549483	
X2	17.9630	7.3581	2.441	0.014670	*
X3	-18.6590	13.4128	-1.391	0.164245	
X4	1.6127	16.9051	0.095	0.924001	
X5	-34.3796	12.0824	-2.845	0.004453	**
X6	26.7523	22.9111	1.168	0.242999	
X7	15.9256	8.7569	1.819	0.069026	.
X8	38.4955	10.8958	3.533	0.000414	***
X9	-15.1075	25.9884	-0.581	0.561054	
X10	21.1043	21.9893	0.960	0.337225	
X11	-15.3352	20.6194	-0.744	0.457076	
X12	-37.2553	30.1766	-1.235	0.217046	
X13	5.1040	4.9030	1.041	0.297932	
X14	-17.9179	18.5317	-0.967	0.333652	
X15	-34.4795	21.4893	-1.604	0.108667	
X16	-16.1786	9.1779	-1.763	0.077999	.
X17	6.6638	7.6930	0.866	0.386413	
X18	22.4296	11.0357	2.032	0.042159	*
X19	-5.5806	62.3680	-0.089	0.928705	
X20	0.4323	27.0876	0.016	0.987267	
X21	24.4507	25.6647	0.953	0.340788	
X22	-19.6151	26.3955	-0.743	0.457441	
X23	205.3235	38.6416	5.314	1.12e-07	***
X24	13.7851	25.9159	0.532	0.594807	
X25	-43.1732	46.0034	-0.938	0.348043	
X26	-18.6249	28.0139	-0.665	0.506180	
X27	3.3282	12.8657	0.259	0.795889	
X28	85.8914	9.6180	8.930	< 2e-16	***
X29	-2.9214	34.2666	-0.085	0.932063	
X30	-20.4975	8.9026	-2.302	0.021352	*
X31	42.2709	59.3313	0.712	0.476215	
X32	-0.6017	19.6785	-0.031	0.975609	
X33	-13.8534	36.1281	-0.383	0.701400	
X34	-0.1976	14.1826	-0.014	0.988883	
X35	-6.2158	20.6602	-0.301	0.763535	
X36	-29.4929	38.7204	-0.762	0.446281	
X37	-7.2401	12.7867	-0.566	0.571269	
X38	7.0066	12.0475	0.582	0.560877	
X39	-35.5406	38.6759	-0.919	0.358174	
X40	32.8678	32.0262	1.026	0.304809	
X41	-11.2205	9.7861	-1.147	0.251611	

X42	-8.1853	61.3184	-0.133	0.893813	
X43	-42.2629	41.0644	-1.029	0.303441	
X44	-2.2952	57.9680	-0.040	0.968418	
X45	1.2543	21.3212	0.059	0.953092	
X46	-3.6707	8.8298	-0.416	0.677631	
X47	-7.1208	35.4093	-0.201	0.840630	
X48	-18.6686	6.1802	-3.021	0.002534	**
X49	2.5385	37.9402	0.067	0.946657	
X50	3.6325	4.7721	0.761	0.446581	
X51	-32.2461	46.0406	-0.700	0.483719	
X52	-1.6332	10.6578	-0.153	0.878214	
X53	17.5245	20.3763	0.860	0.389805	
X54	43.1124	36.3726	1.185	0.235954	
X55	32.0609	9.8971	3.239	0.001205	**
X56	15.7286	16.1379	0.975	0.329788	
X57	-6.7405	24.5456	-0.275	0.783627	
X58	23.4668	17.6431	1.330	0.183551	
X59	-4.3841	27.8898	-0.157	0.875097	
X60	0.5110	4.6539	0.110	0.912564	
X61	98.0082	16.0715	6.098	1.15e-09	***
X62	33.3168	37.8624	0.880	0.378931	
X63	-9.9268	18.0163	-0.551	0.581663	
X64	12.4746	19.0917	0.653	0.513526	
X65	-1.1418	20.3135	-0.056	0.955179	
X66	-42.5283	12.0771	-3.521	0.000433	***
X67	34.4864	57.8463	0.596	0.551087	
X68	-5.5437	21.8095	-0.254	0.799361	
X69	-7.9486	6.9949	-1.136	0.255868	
X70	4.4131	18.5417	0.238	0.811885	
X71	26.3885	14.1471	1.865	0.062198	.
X72	34.9842	18.6426	1.877	0.060634	.
X73	13.9256	33.4201	0.417	0.676927	
X74	-17.2672	12.9866	-1.330	0.183704	
X75	-24.4442	16.6333	-1.470	0.141735	
X76	-15.9993	8.9111	-1.795	0.072643	.
X77	-7.5545	27.5096	-0.275	0.783624	
X78	11.8342	45.8419	0.258	0.796300	
X79	4.6446	51.2620	0.091	0.927810	
X80	-38.1432	6.0465	-6.308	3.06e-10	***
X81	106.2650	77.3569	1.374	0.169596	
X82	-9.6327	32.4234	-0.297	0.766410	
X83	34.4649	17.1630	2.008	0.044687	*
X84	17.2982	28.3991	0.609	0.542478	
X85	-33.1791	13.5693	-2.445	0.014512	*
X86	-10.5915	6.3219	-1.675	0.093923	.
X87	11.2291	18.3558	0.612	0.540733	
X88	22.4524	45.8504	0.490	0.624376	
X89	-223.9878	9.2241	-24.283	< 2e-16	***
X90	22.8047	5.1175	4.456	8.52e-06	***
X91	-15.1646	8.3210	-1.822	0.068446	.
X92	8.2125	5.2855	1.554	0.120302	

X93	8.2922	3.7841	2.191	0.028472	*
X94	-12.6929	5.3516	-2.372	0.017738	*
X95	-6.6693	6.6337	-1.005	0.314768	
X96	19.0848	32.6404	0.585	0.558777	
X97	-35.7708	33.1736	-1.078	0.280955	
X98	-0.5084	18.7473	-0.027	0.978365	
X99	-61.0299	23.5535	-2.591	0.009594	**
X100	3.1035	46.5931	0.067	0.946896	
X101	-0.5405	28.5875	-0.019	0.984917	
X102	-7.1223	28.6834	-0.248	0.803907	
X103	23.9289	47.9917	0.499	0.618079	
X104	22.3633	49.1571	0.455	0.649175	
X105	-15.5601	13.6421	-1.141	0.254092	
X106	17.9068	32.4242	0.552	0.580789	
X107	4.4366	11.7656	0.377	0.706128	
X108	18.6260	10.8078	1.723	0.084880	.
X109	-42.4445	20.9405	-2.027	0.042725	*
X110	9.3726	4.7308	1.981	0.047625	*
X111	17.9210	18.1258	0.989	0.322857	
X112	7.2947	28.7408	0.254	0.799652	
X113	40.2269	73.1782	0.550	0.582541	
X114	-49.4861	23.1965	-2.133	0.032945	*
X115	-28.7995	19.1271	-1.506	0.132209	
X116	-29.3265	7.1976	-4.075	4.68e-05	***
X117	14.0856	66.8071	0.211	0.833021	
X118	-27.6061	16.4682	-1.676	0.093735	.
X119	25.9745	15.5427	1.671	0.094750	.
X120	-0.3702	32.5646	-0.011	0.990931	
X121	-4.9169	6.6240	-0.742	0.457950	
X122	1.3659	4.9094	0.278	0.780846	
X123	25.2357	26.9209	0.937	0.348598	
X124	60.8589	17.6627	3.446	0.000574	***
X125	-14.6320	10.5500	-1.387	0.165528	
X126	6.6508	11.4745	0.580	0.562200	
X127	-11.8409	50.5367	-0.234	0.814759	
X128	-9.9174	16.7614	-0.592	0.554092	
X129	31.1746	57.2788	0.544	0.586286	
X130	0.6182	21.2707	0.029	0.976816	
X131	-170.8985	6.6175	-25.825	< 2e-16	***
X132	-10.6131	60.2022	-0.176	0.860073	
X133	13.0512	40.9510	0.319	0.749966	
X134	31.4181	61.0682	0.514	0.606942	
X135	-21.1232	33.1504	-0.637	0.524028	
X136	12.0545	6.9519	1.734	0.082981	.
X137	58.0809	7.9297	7.324	2.78e-13	***
X138	24.4074	8.2988	2.941	0.003286	**
X139	-36.1491	16.6807	-2.167	0.030272	*
X140	2.0449	9.2516	0.221	0.825079	

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

Residual standard error: 547.9 on 5059 degrees of freedom
Multiple R-squared: 0.9888, Adjusted R-squared: 0.9885
F-statistic: 3202 on 140 and 5059 DF, p-value: < 2.2e-16
# Make predictions.
lm.1.preds <- predict(lm.1, newdata = Data.train)

# Calculate error.
residuals.train <- (lm.1.preds - Data.train$Y)
tibble::tribble(
  ~Metric, ~Value,
  "MSE", mean(residuals.train**2),
  "RMSE", sqrt(mean(residuals.train**2)),
  "MAE", mean(abs(residuals.train))
) %>% print()
# A tibble: 3 x 2
  Metric Value
  <chr> <dbl>
1 MSE 292068.
2 RMSE 540.
3 MAE 383.

```

```
cols = sprintf("X%d", 1:140)
```

```
preprocessor <- caret::preProcess(Data.train %>% select(-Y), method = c("center", "scale"))
```

```
X.train <- predict(preprocessor, Data.train %>% select(-Y)) %>%
  mutate(Y = Data.train$Y)
```

```
X.test <- predict(preprocessor, Data.test)
```

```
summary(X.train)
```

X1		X2		X3		X4		X5	
Min.	:-4.545583	Min.	:-3.45994	Min.	:-3.70724	Min.	:-4.226777	Min.	:-3.4
1st Qu.	:-0.677063	1st Qu.	:-0.68763	1st Qu.	:-0.69468	1st Qu.	:-0.693799	1st Qu.	:-0.6
Median	:-0.001906	Median	: 0.02642	Median	: 0.03088	Median	: 0.001092	Median	: 0.0
Mean	: 0.000000	Mean	: 0.00000	Mean	: 0.00000	Mean	: 0.000000	Mean	: 0.0
3rd Qu.	: 0.688505	3rd Qu.	: 0.70565	3rd Qu.	: 0.70808	3rd Qu.	: 0.688878	3rd Qu.	: 0.6
Max.	: 3.741187	Max.	: 3.62053	Max.	: 2.95962	Max.	: 3.417086	Max.	: 3.2

X7		X8		X9		X10		X11	
Min.	:-3.70696	Min.	:-3.434946	Min.	:-3.25853	Min.	:-3.44871	Min.	:-3.78
1st Qu.	:-0.67428	1st Qu.	:-0.682798	1st Qu.	:-0.70653	1st Qu.	:-0.65495	1st Qu.	:-0.68
Median	: 0.03959	Median	:-0.001462	Median	: 0.02978	Median	: 0.02099	Median	: 0.00
Mean	: 0.00000	Mean	: 0.000000	Mean	: 0.00000	Mean	: 0.00000	Mean	: 0.00
3rd Qu.	: 0.68686	3rd Qu.	: 0.672471	3rd Qu.	: 0.70400	3rd Qu.	: 0.67075	3rd Qu.	: 0.69
Max.	: 3.02927	Max.	: 3.303608	Max.	: 3.24077	Max.	: 3.26056	Max.	: 3.06

X13		X14		X15		X16		X17	
Min.	:-3.56150	Min.	:-3.72314	Min.	:-3.6878	Min.	:-3.97817	Min.	:-3.7018
1st Qu.	:-0.68390	1st Qu.	:-0.67439	1st Qu.	:-0.6935	1st Qu.	:-0.68821	1st Qu.	:-0.6789
Median	: 0.02892	Median	: 0.09632	Median	: 0.0442	Median	: 0.04124	Median	:-0.0010
Mean	: 0.00000	Mean	: 0.00000	Mean	: 0.0000	Mean	: 0.00000	Mean	: 0.0000
3rd Qu.	: 0.71552	3rd Qu.	: 0.72709	3rd Qu.	: 0.7020	3rd Qu.	: 0.69176	3rd Qu.	: 0.6936
Max.	: 3.05694	Max.	: 2.99308	Max.	: 3.0856	Max.	: 3.51637	Max.	: 3.1994

X19	X20	X21	X22	X23
Min. :-3.62900	Min. :-3.41728	Min. :-4.44448	Min. :-3.57852	Min. :-4.27800
1st Qu.: -0.66408	1st Qu.: -0.68472	1st Qu.: -0.67915	1st Qu.: -0.69375	1st Qu.: -0.69400
Median : 0.08493	Median : 0.02317	Median : 0.00635	Median : 0.04158	Median : 0.02900
Mean : 0.00000	Mean : 0.00000	Mean : 0.00000	Mean : 0.00000	Mean : 0.00000
3rd Qu.: 0.70671	3rd Qu.: 0.70330	3rd Qu.: 0.68595	3rd Qu.: 0.70982	3rd Qu.: 0.70600
Max. : 3.11667	Max. : 3.31678	Max. : 3.15069	Max. : 3.47824	Max. : 3.30700
X25	X26	X27	X28	X29
Min. :-3.73624	Min. :-3.57180	Min. :-3.80557	Min. :-4.03721	Min. :-3.38934
1st Qu.: -0.66968	1st Qu.: -0.68330	1st Qu.: -0.66692	1st Qu.: -0.67782	1st Qu.: -0.66390
Median : 0.02494	Median : 0.02180	Median : 0.03583	Median : 0.02485	Median : 0.03730
Mean : 0.00000	Mean : 0.00000	Mean : 0.00000	Mean : 0.00000	Mean : 0.00000
3rd Qu.: 0.69176	3rd Qu.: 0.68780	3rd Qu.: 0.71079	3rd Qu.: 0.69544	3rd Qu.: 0.70120
Max. : 3.19532	Max. : 3.36040	Max. : 3.27093	Max. : 3.08240	Max. : 3.16830
X31	X32	X33	X34	X35
Min. :-3.69673	Min. :-3.59631	Min. :-4.10738	Min. :-3.31469	Min. :-3.31400
1st Qu.: -0.69309	1st Qu.: -0.66419	1st Qu.: -0.69479	1st Qu.: -0.68199	1st Qu.: -0.69000
Median : 0.03219	Median : 0.03618	Median : 0.02613	Median : 0.00505	Median : 0.01400
Mean : 0.00000	Mean : 0.00000	Mean : 0.00000	Mean : 0.00000	Mean : 0.00000
3rd Qu.: 0.69353	3rd Qu.: 0.69227	3rd Qu.: 0.70162	3rd Qu.: 0.67636	3rd Qu.: 0.72700
Max. : 3.31153	Max. : 3.00975	Max. : 3.23025	Max. : 3.28173	Max. : 3.20500
X37	X38	X39	X40	X41
Min. :-3.79650	Min. :-3.55565	Min. :-3.69950	Min. :-3.34399	Min. :-3.61700
1st Qu.: -0.67030	1st Qu.: -0.68884	1st Qu.: -0.67151	1st Qu.: -0.65922	1st Qu.: -0.71100
Median : 0.01277	Median : 0.02428	Median : 0.01987	Median : 0.04424	Median : 0.01700
Mean : 0.00000	Mean : 0.00000	Mean : 0.00000	Mean : 0.00000	Mean : 0.00000
3rd Qu.: 0.68045	3rd Qu.: 0.71075	3rd Qu.: 0.69399	3rd Qu.: 0.70466	3rd Qu.: 0.72600
Max. : 3.42649	Max. : 3.28551	Max. : 3.58447	Max. : 3.68146	Max. : 3.47300
X43	X44	X45	X46	X47
Min. :-3.84705	Min. :-4.32041	Min. :-4.11743	Min. :-3.70779	Min. :-4.61200
1st Qu.: -0.68836	1st Qu.: -0.66949	1st Qu.: -0.67022	1st Qu.: -0.66265	1st Qu.: -0.67900
Median : 0.07077	Median : 0.03669	Median : 0.01368	Median : 0.05286	Median : 0.03500
Mean : 0.00000	Mean : 0.00000	Mean : 0.00000	Mean : 0.00000	Mean : 0.00000
3rd Qu.: 0.72156	3rd Qu.: 0.70789	3rd Qu.: 0.67796	3rd Qu.: 0.69332	3rd Qu.: 0.70000
Max. : 3.36600	Max. : 3.13123	Max. : 3.21373	Max. : 3.21232	Max. : 3.12300
X49	X50	X51	X52	X53
Min. :-4.31367	Min. :-4.20989	Min. :-3.81069	Min. :-3.618123	Min. :-3.88000
1st Qu.: -0.67606	1st Qu.: -0.68277	1st Qu.: -0.69136	1st Qu.: -0.675906	1st Qu.: -0.67000
Median : 0.03681	Median : 0.04206	Median : 0.02203	Median : -0.006447	Median : 0.02000
Mean : 0.00000	Mean : 0.00000	Mean : 0.00000	Mean : 0.000000	Mean : 0.00000
3rd Qu.: 0.70105	3rd Qu.: 0.69888	3rd Qu.: 0.70052	3rd Qu.: 0.674906	3rd Qu.: 0.70000
Max. : 3.20564	Max. : 3.71127	Max. : 3.56923	Max. : 3.733324	Max. : 3.35000
X55	X56	X57	X58	X59
Min. :-3.67391	Min. :-3.74205	Min. :-3.801600	Min. :-3.33475	Min. :-3.45000
1st Qu.: -0.67239	1st Qu.: -0.67386	1st Qu.: -0.680749	1st Qu.: -0.67694	1st Qu.: -0.69000
Median : 0.04701	Median : 0.02848	Median : -0.002649	Median : 0.01399	Median : 0.00000
Mean : 0.00000	Mean : 0.00000	Mean : 0.000000	Mean : 0.00000	Mean : 0.00000
3rd Qu.: 0.70492	3rd Qu.: 0.68661	3rd Qu.: 0.678177	3rd Qu.: 0.69319	3rd Qu.: 0.69000
Max. : 3.26589	Max. : 3.82734	Max. : 3.656320	Max. : 3.28335	Max. : 3.31000
X61	X62	X63	X64	X65
Min. :-4.09989	Min. :-3.490769	Min. :-3.41429	Min. :-4.34194	Min. :-3.60000

1st Qu.: -0.67012	1st Qu.: -0.698439	1st Qu.: -0.67096	1st Qu.: -0.66416	1st Qu.: -0.677
Median : 0.03798	Median : 0.005083	Median : 0.02328	Median : 0.02351	Median : 0.02
Mean : 0.00000	Mean : 0.000000	Mean : 0.00000	Mean : 0.00000	Mean : 0.00
3rd Qu.: 0.70242	3rd Qu.: 0.701811	3rd Qu.: 0.69488	3rd Qu.: 0.69831	3rd Qu.: 0.69
Max. : 3.10769	Max. : 3.131244	Max. : 3.50220	Max. : 3.18449	Max. : 3.44
X67	X68	X69	X70	X71
Min. : -3.69317	Min. : -3.45872	Min. : -4.101076	Min. : -3.84396	Min. : -3.06
1st Qu.: -0.68844	1st Qu.: -0.70903	1st Qu.: -0.678698	1st Qu.: -0.68083	1st Qu.: -0.70
Median : 0.05775	Median : 0.05349	Median : -0.005633	Median : 0.06643	Median : 0.00
Mean : 0.00000	Mean : 0.00000	Mean : 0.000000	Mean : 0.00000	Mean : 0.00
3rd Qu.: 0.71196	3rd Qu.: 0.70514	3rd Qu.: 0.705544	3rd Qu.: 0.71631	3rd Qu.: 0.69
Max. : 3.26973	Max. : 3.04680	Max. : 3.470512	Max. : 2.98337	Max. : 3.00
X73	X74	X75	X76	X77
Min. : -3.73003	Min. : -3.91742	Min. : -4.82517	Min. : -3.58840	Min. : -3.488
1st Qu.: -0.66679	1st Qu.: -0.68724	1st Qu.: -0.66587	1st Qu.: -0.68550	1st Qu.: -0.664
Median : 0.01179	Median : 0.03393	Median : 0.02056	Median : 0.00871	Median : 0.036
Mean : 0.00000	Mean : 0.00000	Mean : 0.00000	Mean : 0.00000	Mean : 0.000
3rd Qu.: 0.69005	3rd Qu.: 0.71884	3rd Qu.: 0.69189	3rd Qu.: 0.69632	3rd Qu.: 0.707
Max. : 3.66479	Max. : 3.40585	Max. : 3.87869	Max. : 3.60109	Max. : 3.113
X79	X80	X81	X82	X83
Min. : -3.79422	Min. : -3.63201	Min. : -3.63162	Min. : -3.54414	Min. : -3.759
1st Qu.: -0.68432	1st Qu.: -0.67902	1st Qu.: -0.69737	1st Qu.: -0.67331	1st Qu.: -0.703
Median : 0.03689	Median : 0.01843	Median : 0.07173	Median : 0.03159	Median : 0.049
Mean : 0.00000	Mean : 0.00000	Mean : 0.00000	Mean : 0.00000	Mean : 0.000
3rd Qu.: 0.70351	3rd Qu.: 0.70762	3rd Qu.: 0.72127	3rd Qu.: 0.69971	3rd Qu.: 0.719
Max. : 3.04386	Max. : 3.32931	Max. : 2.98007	Max. : 3.44648	Max. : 2.912
X85	X86	X87	X88	X89
Min. : -3.48128	Min. : -3.68289	Min. : -3.46947	Min. : -3.48401	Min. : -3.528
1st Qu.: -0.69046	1st Qu.: -0.66690	1st Qu.: -0.67779	1st Qu.: -0.65893	1st Qu.: -0.694
Median : 0.05842	Median : 0.04347	Median : 0.06971	Median : 0.04453	Median : -0.002
Mean : 0.00000	Mean : 0.00000	Mean : 0.00000	Mean : 0.00000	Mean : 0.000
3rd Qu.: 0.70686	3rd Qu.: 0.66710	3rd Qu.: 0.71202	3rd Qu.: 0.68890	3rd Qu.: 0.682
Max. : 2.98515	Max. : 2.91118	Max. : 3.01427	Max. : 3.14931	Max. : 3.449
X91	X92	X93	X94	X95
Min. : -3.5402	Min. : -3.734050	Min. : -3.59734	Min. : -3.33643	Min. : -3.448
1st Qu.: -0.6960	1st Qu.: -0.668128	1st Qu.: -0.67103	1st Qu.: -0.67988	1st Qu.: -0.694
Median : 0.0389	Median : 0.001655	Median : 0.02861	Median : -0.01864	Median : 0.045
Mean : 0.0000	Mean : 0.000000	Mean : 0.00000	Mean : 0.00000	Mean : 0.000
3rd Qu.: 0.7184	3rd Qu.: 0.682239	3rd Qu.: 0.71049	3rd Qu.: 0.68340	3rd Qu.: 0.729
Max. : 3.0649	Max. : 3.417558	Max. : 3.27375	Max. : 3.42103	Max. : 2.919
X97	X98	X99	X100	X101
Min. : -3.969137	Min. : -3.39953	Min. : -3.40050	Min. : -3.459925	Min. : -4.7
1st Qu.: -0.669275	1st Qu.: -0.70527	1st Qu.: -0.67490	1st Qu.: -0.695626	1st Qu.: -0.6
Median : -0.000354	Median : 0.01945	Median : 0.01576	Median : 0.005352	Median : 0.0
Mean : 0.000000	Mean : 0.00000	Mean : 0.00000	Mean : 0.000000	Mean : 0.0
3rd Qu.: 0.689741	3rd Qu.: 0.71678	3rd Qu.: 0.69866	3rd Qu.: 0.709795	3rd Qu.: 0.7
Max. : 3.622854	Max. : 3.14531	Max. : 3.42457	Max. : 3.115759	Max. : 3.6
X103	X104	X105	X106	X107
Min. : -4.35772	Min. : -3.65976	Min. : -3.57156	Min. : -3.86414	Min. : -3.457
1st Qu.: -0.66836	1st Qu.: -0.67562	1st Qu.: -0.69426	1st Qu.: -0.67319	1st Qu.: -0.710
Median : 0.02009	Median : 0.01711	Median : 0.04818	Median : 0.02176	Median : 0.035

Mean : 0.00000	Mean : 0.00000	Mean : 0.00000	Mean : 0.00000	Mean : 0.00000
3rd Qu.: 0.67956	3rd Qu.: 0.69563	3rd Qu.: 0.70696	3rd Qu.: 0.68543	3rd Qu.: 0.71100
Max. : 3.29844	Max. : 3.10512	Max. : 2.91917	Max. : 3.70568	Max. : 3.48300
X109	X110	X111	X112	X113
Min. : -3.8408	Min. : -3.61117	Min. : -3.64743	Min. : -3.60792	Min. : -3.56930
1st Qu.: -0.6703	1st Qu.: -0.63839	1st Qu.: -0.67400	1st Qu.: -0.66748	1st Qu.: -0.70150
Median : 0.0296	Median : 0.02982	Median : 0.01152	Median : 0.01523	Median : 0.02950
Mean : 0.0000	Mean : 0.00000	Mean : 0.00000	Mean : 0.00000	Mean : 0.00000
3rd Qu.: 0.7006	3rd Qu.: 0.69527	3rd Qu.: 0.67470	3rd Qu.: 0.68031	3rd Qu.: 0.69140
Max. : 3.2327	Max. : 3.22039	Max. : 3.18653	Max. : 2.98220	Max. : 3.46580
X115	X116	X117	X118	X119
Min. : -4.31972	Min. : -4.03922	Min. : -4.368147	Min. : -4.12963	Min. : -3.84000
1st Qu.: -0.68514	1st Qu.: -0.68241	1st Qu.: -0.667562	1st Qu.: -0.69043	1st Qu.: -0.67000
Median : 0.05283	Median : 0.04655	Median : 0.002564	Median : 0.02468	Median : 0.02000
Mean : 0.00000	Mean : 0.00000	Mean : 0.000000	Mean : 0.00000	Mean : 0.00000
3rd Qu.: 0.70799	3rd Qu.: 0.70213	3rd Qu.: 0.680001	3rd Qu.: 0.69293	3rd Qu.: 0.68000
Max. : 3.43520	Max. : 3.28030	Max. : 3.571775	Max. : 3.14484	Max. : 3.45000
X121	X122	X123	X124	X125
Min. : -3.59227	Min. : -3.62655	Min. : -3.85509	Min. : -4.27828	Min. : -3.93900
1st Qu.: -0.68082	1st Qu.: -0.69296	1st Qu.: -0.67978	1st Qu.: -0.66257	1st Qu.: -0.69900
Median : -0.01068	Median : 0.03096	Median : 0.03433	Median : 0.01077	Median : 0.03400
Mean : 0.00000	Mean : 0.00000	Mean : 0.00000	Mean : 0.00000	Mean : 0.00000
3rd Qu.: 0.67956	3rd Qu.: 0.69427	3rd Qu.: 0.70168	3rd Qu.: 0.68349	3rd Qu.: 0.68000
Max. : 3.52269	Max. : 3.65959	Max. : 3.19042	Max. : 3.66290	Max. : 2.99800
X127	X128	X129	X130	X131
Min. : -3.9927	Min. : -3.918435	Min. : -3.72954	Min. : -3.68846	Min. : -3.96000
1st Qu.: -0.6774	1st Qu.: -0.691602	1st Qu.: -0.67964	1st Qu.: -0.67392	1st Qu.: -0.67700
Median : 0.0242	Median : 0.000055	Median : 0.03281	Median : 0.01039	Median : 0.01400
Mean : 0.0000	Mean : 0.000000	Mean : 0.00000	Mean : 0.00000	Mean : 0.00000
3rd Qu.: 0.6975	3rd Qu.: 0.696643	3rd Qu.: 0.72118	3rd Qu.: 0.68752	3rd Qu.: 0.68700
Max. : 3.1184	Max. : 3.148087	Max. : 2.96918	Max. : 2.98133	Max. : 3.72600
X133	X134	X135	X136	X137
Min. : -3.69899	Min. : -3.3438	Min. : -3.403281	Min. : -3.68947	Min. : -4.65800
1st Qu.: -0.70255	1st Qu.: -0.6776	1st Qu.: -0.674726	1st Qu.: -0.68439	1st Qu.: -0.67300
Median : 0.04983	Median : 0.0494	Median : 0.009819	Median : 0.02887	Median : 0.00800
Mean : 0.00000	Mean : 0.0000	Mean : 0.000000	Mean : 0.00000	Mean : 0.00000
3rd Qu.: 0.69953	3rd Qu.: 0.6914	3rd Qu.: 0.705494	3rd Qu.: 0.69136	3rd Qu.: 0.67700
Max. : 3.27543	Max. : 3.2369	Max. : 3.432824	Max. : 3.54074	Max. : 3.30300
X139	X140	Y		
Min. : -3.71590	Min. : -4.08938	Min. : -7028		
1st Qu.: -0.69713	1st Qu.: -0.68850	1st Qu.: 5574		
Median : 0.02157	Median : 0.04886	Median : 9148		
Mean : 0.00000	Mean : 0.00000	Mean : 9234		
3rd Qu.: 0.68861	3rd Qu.: 0.70132	3rd Qu.: 12727		
Max. : 3.43156	Max. : 3.12703	Max. : 31402		

```
lm.2 <- lm(Y ~ ., data = X.train)
print(lm.2)
```

Call:

```
lm(formula = Y ~ ., data = X.train)
```

Coefficients:

(Intercept)	X1	X2	X3	X4	X5	X6
9234.383	-223.607	485.087	-476.938	39.960	-859.334	670.611
X9	X10	X11	X12	X13	X14	X15
-393.001	489.680	-384.870	-1021.144	128.275	-553.832	-927.443
X18	X19	X20	X21	X22	X23	X24
635.292	-148.861	11.176	547.561	-504.959	5325.299	328.696
X27	X28	X29	X30	X31	X32	X33
81.898	2028.659	-81.538	-565.286	1120.025	-16.081	-325.043
X36	X37	X38	X39	X40	X41	X42
-755.532	-161.325	186.291	-860.705	887.856	-280.115	-221.539
X45	X46	X47	X48	X49	X50	X51
30.482	-94.371	-191.962	-477.175	64.598	95.294	-778.584
X54	X55	X56	X57	X58	X59	X60
1036.022	772.790	417.942	-170.784	602.940	-104.744	14.050
X63	X64	X65	X66	X67	X68	X69
-235.648	306.058	-27.380	-1056.157	916.666	-144.325	-181.970
X72	X73	X74	X75	X76	X77	X78
776.492	342.317	-428.276	-573.296	-425.842	-189.643	286.878
X81	X82	X83	X84	X85	X86	X87
2919.517	-247.263	1048.379	474.052	-923.788	-246.748	303.123
X90	X91	X92	X93	X94	X95	X96
604.926	-416.315	181.762	220.433	-331.117	-209.480	440.192
X99	X100	X101	X102	X103	X104	X105
-1450.076	89.957	-13.600	-193.171	639.373	596.519	-428.933
X108	X109	X110	X111	X112	X113	X114
447.126	-1094.195	236.048	442.460	187.456	977.049	-1259.103
X117	X118	X119	X120	X121	X122	X123
330.827	-671.394	598.704	-10.723	-125.578	32.869	666.234
X126	X127	X128	X129	X130	X131	X132
160.150	-299.069	-253.486	801.497	15.668	-3877.795	-265.574
X135	X136	X137	X138	X139	X140	
-557.545	327.802	1354.143	693.605	-877.879	54.562	

Print summary.

summary(lm.2) %>% print()

Call:

lm(formula = Y ~ ., data = X.train)

Residuals:

Min	1Q	Median	3Q	Max
-2425.8	-297.7	-38.2	247.7	5937.1

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	9234.383	7.598	1215.341	< 2e-16 ***
X1	-223.607	373.566	-0.599	0.549483
X2	485.087	198.703	2.441	0.014670 *
X3	-476.938	342.841	-1.391	0.164245
X4	39.960	418.871	0.095	0.924001
X5	-859.334	302.004	-2.845	0.004453 **

X6	670.611	574.321	1.168	0.242999	
X7	420.164	231.033	1.819	0.069026	.
X8	919.105	260.144	3.533	0.000414	***
X9	-393.001	676.055	-0.581	0.561054	
X10	489.680	510.215	0.960	0.337225	
X11	-384.870	517.489	-0.744	0.457076	
X12	-1021.144	827.122	-1.235	0.217046	
X13	128.275	123.225	1.041	0.297932	
X14	-553.832	572.805	-0.967	0.333652	
X15	-927.443	578.028	-1.604	0.108667	
X16	-407.095	230.939	-1.763	0.077999	.
X17	176.900	204.221	0.866	0.386413	
X18	635.292	312.572	2.032	0.042159	*
X19	-148.861	1663.643	-0.089	0.928705	
X20	11.176	700.277	0.016	0.987267	
X21	547.561	574.747	0.953	0.340788	
X22	-504.959	679.508	-0.743	0.457441	
X23	5325.299	1002.214	5.314	1.12e-07	***
X24	328.696	617.947	0.532	0.594807	
X25	-1087.499	1158.790	-0.938	0.348043	
X26	-461.598	694.293	-0.665	0.506180	
X27	81.898	316.595	0.259	0.795889	
X28	2028.659	227.166	8.930	< 2e-16	***
X29	-81.538	956.408	-0.085	0.932063	
X30	-565.286	245.518	-2.302	0.021352	*
X31	1120.025	1572.061	0.712	0.476215	
X32	-16.081	525.932	-0.031	0.975609	
X33	-325.043	847.673	-0.383	0.701400	
X34	-4.949	355.140	-0.014	0.988883	
X35	-177.239	589.113	-0.301	0.763535	
X36	-755.532	991.915	-0.762	0.446281	
X37	-161.325	284.916	-0.566	0.571269	
X38	186.291	320.319	0.582	0.560877	
X39	-860.705	936.633	-0.919	0.358174	
X40	887.856	865.121	1.026	0.304809	
X41	-280.115	244.306	-1.147	0.251611	
X42	-221.539	1659.620	-0.133	0.893813	
X43	-1146.384	1113.874	-1.029	0.303441	
X44	-61.950	1564.610	-0.040	0.968418	
X45	30.482	518.165	0.059	0.953092	
X46	-94.371	227.005	-0.416	0.677631	
X47	-191.962	954.566	-0.201	0.840630	
X48	-477.175	157.968	-3.021	0.002534	**
X49	64.598	965.458	0.067	0.946657	
X50	95.294	125.192	0.761	0.446581	
X51	-778.584	1111.652	-0.700	0.483719	
X52	-41.506	270.854	-0.153	0.878214	
X53	442.492	514.498	0.860	0.389805	
X54	1036.022	874.058	1.185	0.235954	
X55	772.790	238.558	3.239	0.001205	**
X56	417.942	428.819	0.975	0.329788	

X57	-170.784	621.914	-0.275	0.783627	
X58	602.940	453.310	1.330	0.183551	
X59	-104.744	666.330	-0.157	0.875097	
X60	14.050	127.946	0.110	0.912564	
X61	2565.091	420.628	6.098	1.15e-09	***
X62	871.923	990.883	0.880	0.378931	
X63	-235.648	427.679	-0.551	0.581663	
X64	306.058	468.405	0.653	0.513526	
X65	-27.380	487.129	-0.056	0.955179	
X66	-1056.157	299.925	-3.521	0.000433	***
X67	916.666	1537.584	0.596	0.551087	
X68	-144.325	567.791	-0.254	0.799361	
X69	-181.970	160.137	-1.136	0.255868	
X70	120.757	507.368	0.238	0.811885	
X71	692.198	371.094	1.865	0.062198	.
X72	776.492	413.782	1.877	0.060634	.
X73	342.317	821.525	0.417	0.676927	
X74	-428.276	322.105	-1.330	0.183704	
X75	-573.296	390.106	-1.470	0.141735	
X76	-425.842	237.179	-1.795	0.072643	.
X77	-189.643	690.581	-0.275	0.783624	
X78	286.878	1111.278	0.258	0.796300	
X79	121.238	1338.079	0.091	0.927810	
X80	-1021.427	161.917	-6.308	3.06e-10	***
X81	2919.517	2125.298	1.374	0.169596	
X82	-247.263	832.282	-0.297	0.766410	
X83	1048.379	522.077	2.008	0.044687	*
X84	474.052	778.269	0.609	0.542478	
X85	-923.788	377.802	-2.445	0.014512	*
X86	-246.748	147.279	-1.675	0.093923	.
X87	303.123	495.504	0.612	0.540733	
X88	542.819	1108.499	0.490	0.624376	
X89	-5701.825	234.809	-24.283	< 2e-16	***
X90	604.926	135.750	4.456	8.52e-06	***
X91	-416.315	228.437	-1.822	0.068446	.
X92	181.762	116.981	1.554	0.120302	
X93	220.433	100.593	2.191	0.028472	*
X94	-331.117	139.605	-2.372	0.017738	*
X95	-209.480	208.362	-1.005	0.314768	
X96	440.192	752.853	0.585	0.558777	
X97	-933.925	866.115	-1.078	0.280955	
X98	-13.898	512.444	-0.027	0.978365	
X99	-1450.076	559.634	-2.591	0.009594	**
X100	89.957	1350.542	0.067	0.946896	
X101	-13.600	719.375	-0.019	0.984917	
X102	-193.171	777.951	-0.248	0.803907	
X103	639.373	1282.324	0.499	0.618079	
X104	596.519	1311.217	0.455	0.649175	
X105	-428.933	376.060	-1.141	0.254092	
X106	445.517	806.705	0.552	0.580789	
X107	110.537	293.138	0.377	0.706128	

```

X108      447.126      259.447      1.723 0.084880 .
X109     -1094.195      539.836     -2.027 0.042725 *
X110       236.048      119.145      1.981 0.047625 *
X111       442.460      447.517      0.989 0.322857
X112       187.456      738.565      0.254 0.799652
X113       977.049     1777.383      0.550 0.582541
X114     -1259.103      590.202     -2.133 0.032945 *
X115     -791.420      525.619     -1.506 0.132209
X116     -770.744      189.163     -4.075 4.68e-05 ***
X117       330.827     1569.091      0.211 0.833021
X118     -671.394      400.514     -1.676 0.093735 .
X119       598.704      358.255      1.671 0.094750 .
X120      -10.723      943.357     -0.011 0.990931
X121     -125.578      169.177     -0.742 0.457950
X122        32.869      118.136      0.278 0.780846
X123       666.234      710.726      0.937 0.348598
X124      1437.531      417.206      3.446 0.000574 ***
X125     -396.275      285.723     -1.387 0.165528
X126       160.150      276.305      0.580 0.562200
X127     -299.069     1276.415     -0.234 0.814759
X128     -253.486      428.418     -0.592 0.554092
X129       801.497     1472.633      0.544 0.586286
X130        15.668      539.102      0.029 0.976816
X131    -3877.795      150.154    -25.825 < 2e-16 ***
X132     -265.574     1506.455     -0.176 0.860073
X133       363.296     1139.924      0.319 0.749966
X134       809.874     1574.173      0.514 0.606942
X135     -557.545      875.001     -0.637 0.524028
X136       327.802      189.045      1.734 0.082981 .
X137      1354.143      184.879      7.324 2.78e-13 ***
X138       693.605      235.834      2.941 0.003286 **
X139     -877.879      405.089     -2.167 0.030272 *
X140        54.562      246.856      0.221 0.825079
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 547.9 on 5059 degrees of freedom
Multiple R-squared:  0.9888,    Adjusted R-squared:  0.9885
F-statistic: 3202 on 140 and 5059 DF,  p-value: < 2.2e-16
# Make predictions.
lm.2.preds <- predict(lm.2, newdata = X.train)

# Calculate error.
residuals.train <- (lm.2.preds - X.train$Y)
tibble::tribble(
  ~Metric, ~Value,
  "MSE", mean(residuals.train**2),
  "RMSE", sqrt(mean(residuals.train**2)),
  "MAE", mean(abs(residuals.train))
) %>% print()
# A tibble: 3 x 2

```

	Metric	Value
	<chr>	<dbl>
1	MSE	292068.
2	RMSE	540.
3	MAE	383.

```
rf.1 <- randomForest::randomForest(Y ~ ., data = X.train)
print(rf.1)
```

Call:

```
randomForest(formula = Y ~ ., data = X.train)
      Type of random forest: regression
      Number of trees: 500
No. of variables tried at each split: 46
```

```
      Mean of squared residuals: 5706318
      % Var explained: 78.2
```

Print summary.

```
summary(rf.1) %>% print()
```

	Length	Class	Mode
call	3	-none-	call
type	1	-none-	character
predicted	5200	-none-	numeric
mse	500	-none-	numeric
rsq	500	-none-	numeric
oob.times	5200	-none-	numeric
importance	140	-none-	numeric
importanceSD	0	-none-	NULL
localImportance	0	-none-	NULL
proximity	0	-none-	NULL
ntree	1	-none-	numeric
mtry	1	-none-	numeric
forest	11	-none-	list
coefs	0	-none-	NULL
y	5200	-none-	numeric
test	0	-none-	NULL
inbag	0	-none-	NULL
terms	3	terms	call

Make predictions.

```
rf.1.preds <- predict(rf.1, newdata = X.train)
```

Calculate error.

```
residuals.train <- (rf.1.preds - X.train$Y)
```

```
tibble::tribble(
  ~Metric, ~Value,
  "MSE", mean(residuals.train**2),
  "RMSE", sqrt(mean(residuals.train**2)),
  "MAE", mean(abs(residuals.train))
) %>% print()
```

A tibble: 3 x 2

Metric	Value
--------	-------

<chr>	<dbl>
1 MSE	874098.
2 RMSE	935.
3 MAE	720.

randomForest::importance(rf.1)	
	IncNodePurity
X1	279706877
X2	295617503
X3	231800590
X4	315459510
X5	250783247
X6	5389227047
X7	266283389
X8	297157110
X9	415154842
X10	321120185
X11	242502823
X12	195112194
X13	231548368
X14	5823043891
X15	345233159
X16	236239461
X17	291637694
X18	237699511
X19	2482086182
X20	218778234
X21	241381353
X22	378327338
X23	28958832808
X24	235021854
X25	263074427
X26	356209512
X27	242198849
X28	421811892
X29	231124826
X30	683755213
X31	216436039
X32	221752581
X33	239691711
X34	229434604
X35	278655229
X36	222490426
X37	652468561
X38	268085704
X39	254886719
X40	247345666
X41	265861217
X42	319488745
X43	271391355
X44	298901133

X45	244205777
X46	323958106
X47	788974993
X48	310675768
X49	233175502
X50	463831544
X51	236084501
X52	279009849
X53	689417769
X54	233646901
X55	417230212
X56	220305865
X57	284744854
X58	216600213
X59	382250950
X60	266392783
X61	10276562913
X62	240050557
X63	221219607
X64	678147438
X65	311386852
X66	311555783
X67	375895853
X68	366570227
X69	220270222
X70	509899284
X71	299010393
X72	255668522
X73	255065942
X74	1005719865
X75	506771095
X76	617420461
X77	301382390
X78	276116695
X79	262283291
X80	265705319
X81	12485669351
X82	213277045
X83	271998027
X84	278041636
X85	320017132
X86	255892163
X87	220562530
X88	218917832
X89	17924246250
X90	213218493
X91	247246992
X92	379330216
X93	242928043
X94	270228240
X95	368519929

X96	486742614
X97	247253906
X98	201428692
X99	704740359
X100	228391694
X101	242653803
X102	335060546
X103	242488851
X104	256253574
X105	234020913
X106	297649271
X107	284020915
X108	259893885
X109	653954544
X110	275534365
X111	295739482
X112	285403994
X113	311850813
X114	274583146
X115	234929730
X116	299233678
X117	316306969
X118	290686161
X119	404904147
X120	230486023
X121	238441243
X122	245087208
X123	308902766
X124	321695757
X125	241547947
X126	266584442
X127	1690824816
X128	298684042
X129	1238540609
X130	251446533
X131	4076649280
X132	270111973
X133	272216528
X134	425613372
X135	596811542
X136	262938884
X137	306164305
X138	230754680
X139	220589630
X140	3902366083

randomForest::varImpPlot(rf.1)

rf.1

