Jet Propulsion Laboratory
California Institute of Technology

Possible Asteroid Impacts with Earth

Tatiksha Singh

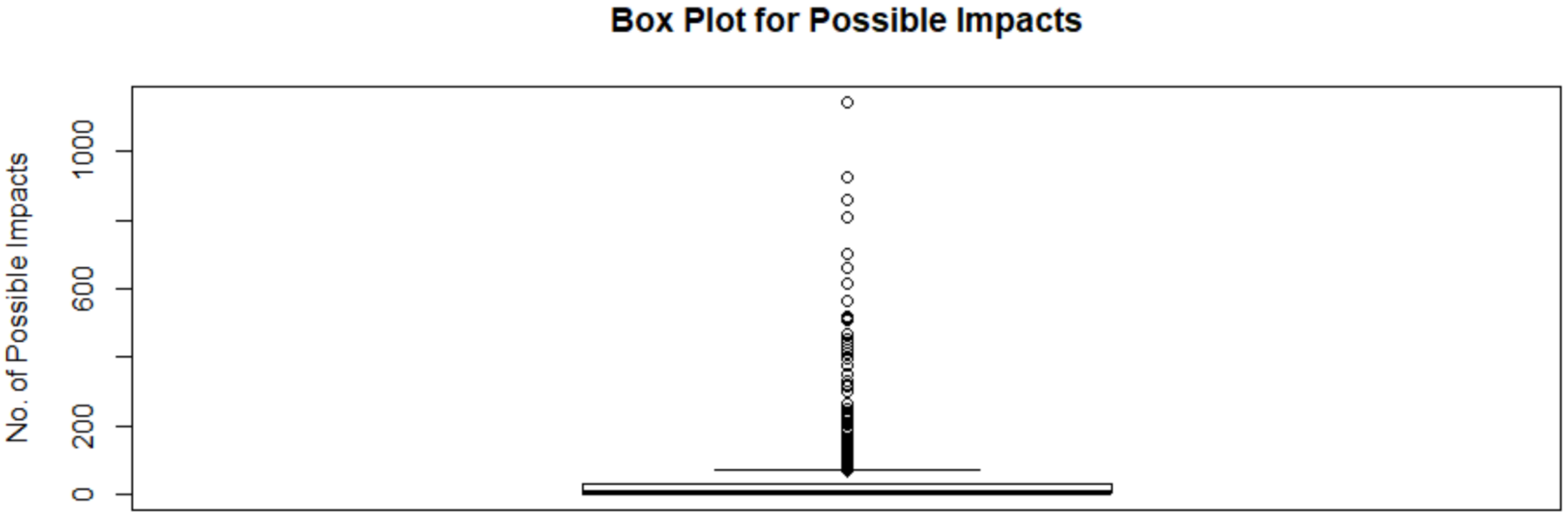
BANA 200

Professor Richard W. Selby

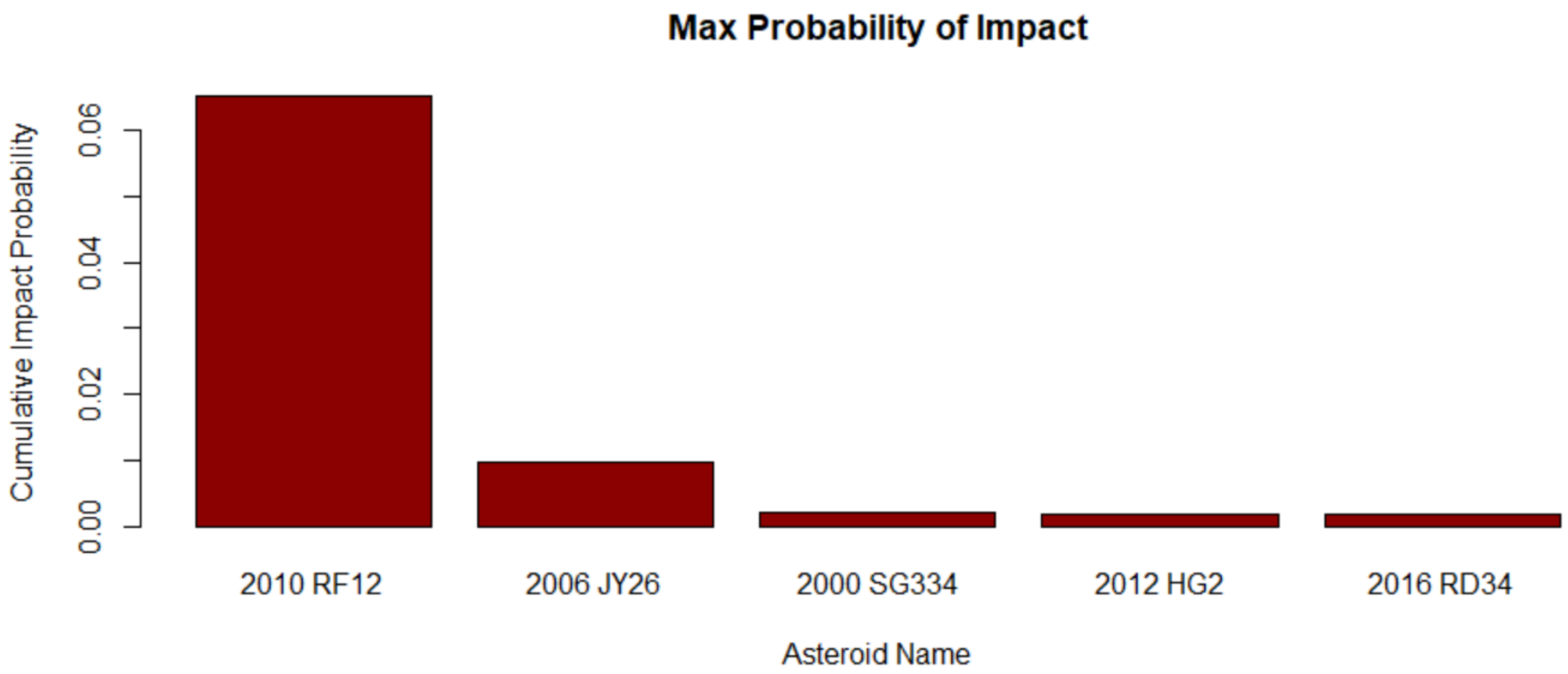
Goal of Asteroid Analysis

- What is the relationship between the number of possible impacts and various variables such as Asteroid Diameter, Velocity, Magnitude, and Impact Probability? Moreover, what asteroids are likely to impact Earth?
- Dependent Variable=
  - Possible Impacts
- Independent Variables=
  - Cumulative Impact Probability
  - Asteroid Velocity (km/s)*
  - Asteroid Magnitude**
  - Asteroid Diameter (km)
  - Cumulative Palermo***
  - Maximum Palermo****

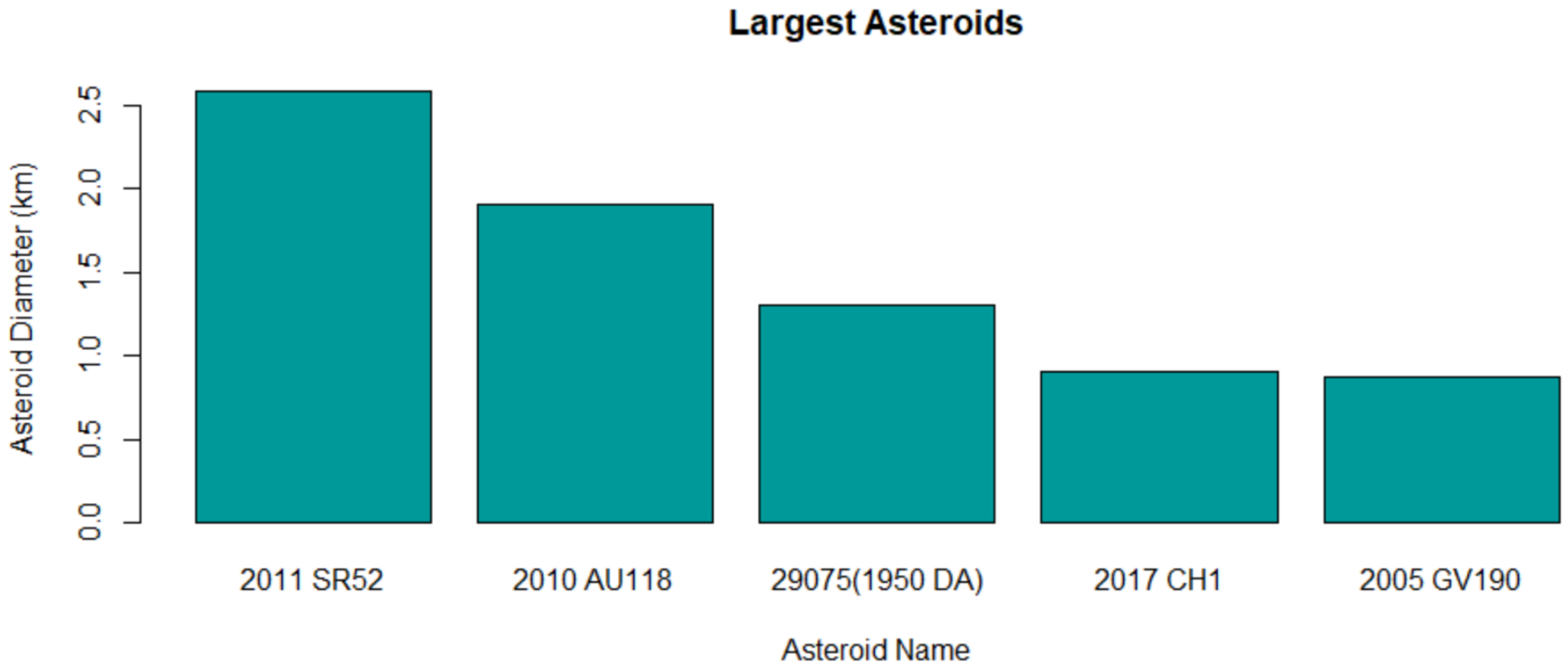
Possible Asteroid Impacts



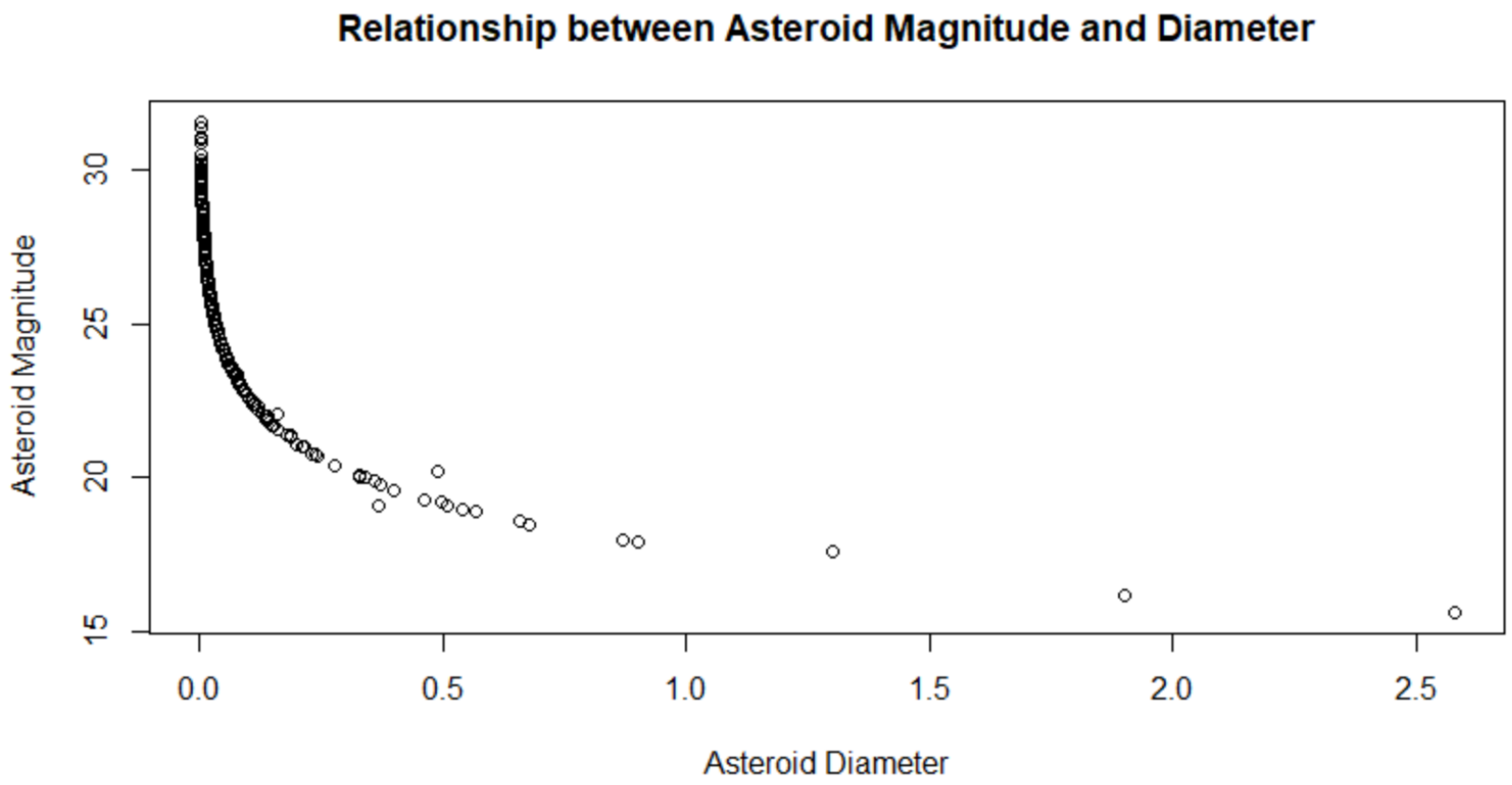
Top 5 asteroids with highest impact probability



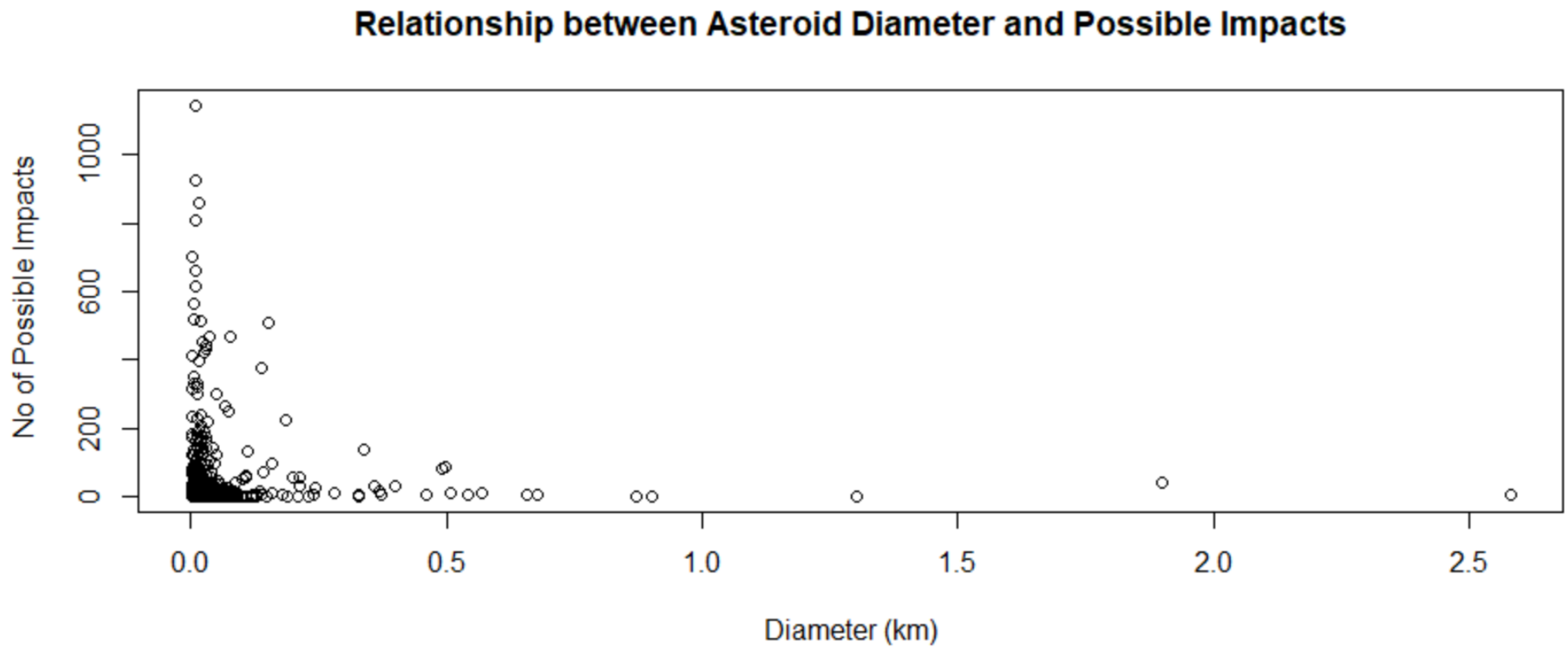
Top 5 Largest Asteroids



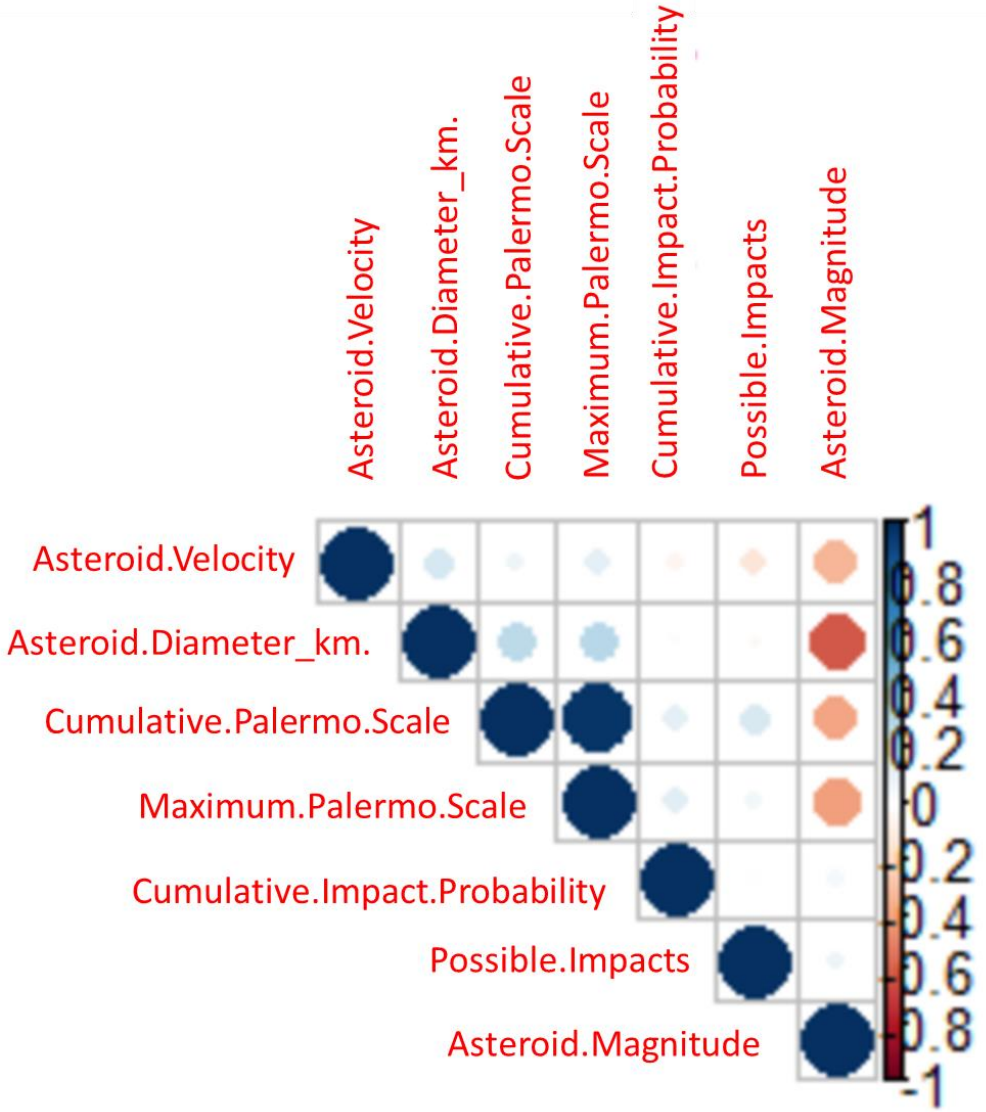
How does Diameter affect Magnitude



How does Diameter affect Possible Impacts



Correlation Matrix for all Variables



Descriptive Stats

Possible.Impacts	
nbr.val	683
nbr.null	0
nbr.na	0
min	1
max	1144
range	1143
sum	29607
median	6
mean	43.348463
SE.mean	4.353414
CI.mean.0.9	8.547703
var	12944.35933
std.dev	113.77328
coef.var	2.624621
Cumulative.Impact.Probability	
nbr.val	6.83E+02
nbr.null	0.00E+00
nbr.na	0.00E+00
min	1.10E-10
max	6.50E-02
range	6.50E-02
sum	1.05E-01
median	1.70E-06
mean	1.54E-04
SE.mean	9.64E-05
CI.mean.0.9	1.89E-04
var	6.35E-06
std.dev	2.52E-03
coef.var	1.64E+01

	Asteroid.Velocity
nbr.val	683
nbr.null	0
nbr.na	0
min	0.34
max	39.47
range	39.13
sum	7828.94
median	10.5
mean	11.4625769
SE.mean	0.2321769
CI.mean.0.95	0.4558673
var	36.8178596
std.dev	6.0677722
coef.var	0.529355
□	Asteroid.Magnitude
nbr.val	6.83E+02
nbr.null	0.00E+00
nbr.na	0.00E+00
min	1.56E+01
max	3.16E+01
range	1.60E+01
sum	1.79E+04
median	2.64E+01
mean	2.62E+01
SE.mean	9.14E-02
CI.mean.0.95	1.79E-01
var	5.70E+00
std.dev	2.39E+00
coef.var	9.13E-02

Cumulative.Palermo.Scale	
nbr.val	6.83E+02
nbr.null	0.00E+00
nbr.na	0.00E+00
min	-1.10E+01
max	-1.42E+00
range	9.56E+00
sum	-4.45E+03
median	-6.46E+00
mean	-6.51E+00
SE.mean	5.77E-02
CI.mean.0.95	1.13E-01
var	2.28E+00
std.dev	1.51E+00
coef.var	-2.32E-01
Maximum.Palermo.Scale	
nbr.val	6.83E+02
nbr.null	0.00E+00
nbr.na	0.00E+00
min	-1.10E+01
max	-1.42E+00
range	9.58E+00
sum	-4.65E+03
median	-6.79E+00
mean	-6.80E+00
SE.mean	5.65E-02
CI.mean.0.95	1.11E-01
var	2.18E+00
std.dev	1.48E+00
coef.var	-2.17E-01

Asteroid.Diameter..km.	
nbr.val	6.83E+02
nbr.null	0.00E+00
nbr.na	0.00E+00
min	2.00E-03
max	2.58E+00
range	2.58E+00
sum	3.37E+01
median	1.70E-02
mean	4.94E-02
SE.mean	5.98E-03
CI.mean.0.95	1.18E-02
var	2.45E-02
std.dev	1.56E-01
coef.var	3.17E+00

Pearson's product-moment correlation

- Dependent Variable= Possible Impacts
- Independent Variables=
  - Cumulative Impact Probability
  - Asteroid Velocity
  - Asteroid Magnitude
  - Asteroid Diameter
  - Cumulative Palermo
  - Maximum Palermo

```
data:cum.impact.prob and pos.impact
s
t = 0.20357, df = 681, p-value = 0.
8388
alternative hypothesis: true correl
ation is not equal to 0
95 percent confidence interval:
 -0.06725892  0.08277194
sample estimates:
cor = 0.007800408
```

```
data:asteroid.diam and pos.impacts
t = -0.88977, df = 681, p-value = 0
.3739
alternative hypothesis: true correl
ation is not equal to 0
95 percent confidence interval:
 -0.10881793  0.04104877
sample estimates:
cor =-0.03407614
```

```
data:  asteroid.vel and pos.impacts
t = -3.6051, df = 681, p-value = 0.
0003349
alternative hypothesis: true correl
ation is not equal to 0
95 percent confidence interval:
 -0.20971371 -0.06246805
sample estimates:
cor = -0.1368467
```

```
data:  cum.palermo and pos.impacts
t = 4.4556, df = 681, p-value = 9.7
78e-06
alternative hypothesis: true correl
ation is not equal to 0
95 percent confidence interval:
 0.09447703 0.24029014
sample estimates:
cor= 0.1683041
```

```
data:  asteroid.mag and pos.impacts
t = 1.9325, df = 681, p-value = 0.0
5371
alternative hypothesis: true correl
ation is not equal to 0
95 percent confidence interval:
 -0.001174617  0.148051577
sample estimates:
cor = 0.07385186
```

```
data:  max.palermo and pos.impacts
t = 1.3082, df = 681, p-value = 0.1
912
alternative hypothesis: true correl
ation is not equal to 0
95 percent confidence interval:
 -0.02504568  0.12462025
sample estimates:
cor =0.05006837
```

Multivariate Linear Regression

```
lm(formula = pos.impacts ~ cum.impact.prob + asteroid.vel + asteroid.mag +
    asteroid.diam + cum.palermo + max.palermo)

Residuals:
    Min      1Q  Median      3Q     Max
-150.55  -40.12   -1.01   20.88 1079.49

Coefficients:
                Estimate Std. Error t value Pr(>|t|)
(Intercept)      -63.3638    54.8264  -1.156    0.248
cum.impact.prob  775.0668  1427.0326   0.543    0.587
asteroid.vel      -0.2493     0.6295  -0.396    0.692
asteroid.mag       3.2995     2.0653   1.598    0.111
asteroid.diam      3.0512    28.6440   0.107    0.915
cum.palermo      212.4975    11.6770  18.198   <2e-16 ***
max.palermo     -206.7428    12.0644 -17.137   <2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 92.34 on 676 degrees of freedom
Multiple R-squared:  0.347,    Adjusted R-squared:  0.3412
F-statistic: 59.88 on 6 and 676 DF,  p-value: < 2.2e-16
```

Multiple Linear Regression Equation

Possible Impacts = Intercept + cum.impact.prob(x1) + asteroid.vel(x2) + asteroid.mag(x3) + asteroid.diam(x4) + cum.palermo(x5) + max.palermo (x6)

Possible Impacts = -63.36 + 775.07(x1) + (-0.25(x2)) + 3.30(x3) + 3.05(x4) + 212.50(x5) + (-206.74(x6))

Key Insights

- Asteroid “2010 RF12” has the highest chance of colliding into Earth
  - However, it only has a 6.5% chance of hitting Earth (a total of 52 possible impacts)
  - Moreover, it has a diameter of 0.007KM which is significantly less than the largest asteroid (“2011 SR52” with a diameter of 2.579km)
- The Largest Asteroid “2011 SR52” only has a cumulative probability of 7.60E-10 with 4 possible impacts
  - It had a velocity of 13.55km/s, and while it did not have the highest velocity, based on its size it is still very dangerous
  - However, we do not need to worry *too* much about it colliding into Earth
- Asteroid Diameter and Magnitude had the highest negative correlation (and overall highest correlation) of -0.612
  - As an Asteroid’s Diameter increased, its magnitude (intrinsic brightness) decreased
  - This observation has some significance after researching the relationship between Magnitude and Diameter on NASA’s website. They calculate asteroids’ diameter based on the absolute magnitude .
  - The following pairs of variables had little to some negative correlation:
    - Asteroid Velocity and Magnitude (-0.33)
    - Cumulative Palermo and Asteroid Magnitude (-0.40)
    - Maximum Palermo and Asteroid Magnitude (-0.41)
- Possible Asteroid Impacts has no correlation between any of the independent variables
  - The highest number of possible impacts was 1144 with a range of 1143
  - It had a lot of variability and outliers in its data
  - Cumulative and Maximum Palermo were statistically significant after performing a multiple variable regression analysis

Appendix

Definitions

*Velocity (km/s) = of the asteroid relative to the Earth, assuming a massless Earth.
**Magnitude=Intrinsic Brightness- It is the apparent magnitude of the object when it is 1 au from both the sun and the observer, and at full phase for the observer.
***Cumulative Palermo=Cumulative hazard rating according to the Palermo technical impact hazard scale, based on the tabulated impact date, impact probability and impact energy.
****Max Palermo= Maximum hazard rating according to the Palermo technical impact hazard scale, based on the tabulated impact date, impact probability and impact energy.

Definitions taken from NASA CNEOS Website:
[Chodas, P. Sentry: Earth Impact Monitoring. Retrieved from https://cneos.jpl.nasa.gov/sentry/]

Appendix (code screenshots)

File Edit Code View Plots Session Build Debug Profile Tools Help

Go to file/function | Addins | Project: (None)

impacts

Filter

Cumulative.Palermo.Scale	Maximum.Palermo.Scale	Maximum.Torino.Scale
-8.31	-8.31	0
-6.60	-6.96	0
-6.48	-6.87	0
-6.83	-6.95	0
-3.85	-4.30	0
-5.03	-5.51	0

Showing 1 to 8 of 683 entries, 11 total columns

Console | Terminal | Jobs

```
> corrplot(M, "circle", type="upper", order= "hclust", tl.cex=0.5, tl.srt = 90, tl.pos= "td")
> plot(asteroid.diam, pos.impacts, main="Relationship between Asteroid Diameter and Possible Impacts", xlab="Diameter (km)", ylab=" No of Possible Impacts", pch=21)
>
```

Environment | History | Connections

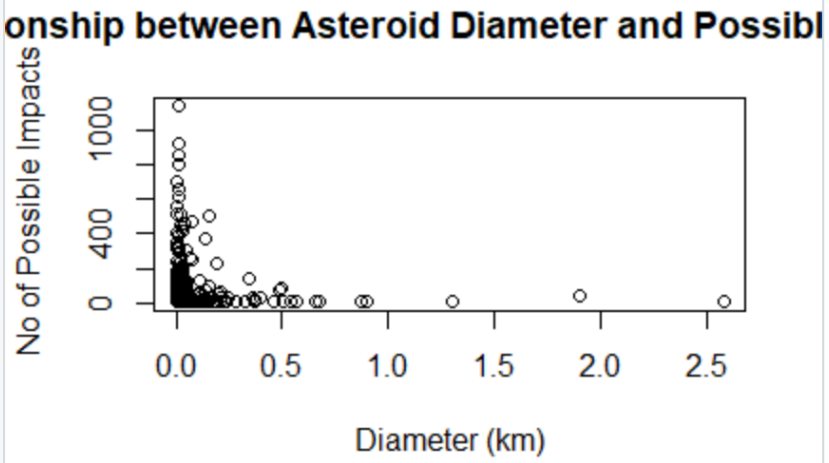
Import Dataset | List

Global Environment

asteroid.mag num [1:683] 28.3 31.4 31.4 26.7 1…

Files | Plots | Packages | Help | Viewer

Zoom | Export



impacts

Filter

Cumulative.Palermo.Scale	Maximum.Palermo.Scale	Maximum.Torino.Scale
-8.31	-8.31	0
-6.60	-6.96	0
-6.48	-6.87	0
-6.83	-6.95	0
-3.85	-4.30	0
-5.03	-5.51	0

Showing 1 to 8 of 683 entries, 11 total columns

Console | Terminal | Jobs

```
ex=0.3, tl.srt = 90, tl.pos= td)
Error in isFALSE(tl.pos) : object 'td' not found
>
> corrplot(M, "circle", type="upper", order= "hclust", tl.c
ex=0.3, tl.srt = 90, tl.pos= "td")
>
```

Environment | History | Connections

Import Dataset | List

Global Environment

```
asteroid.mag  num [1:683] 28.3 31.4 31.4 26.7 1...
```

Files | Plots | Packages | Help | Viewer

Zoom | Export

Console | Terminal | Jobs

~/

```
> stat.desc(impacts[, c(4:11)], p=0.95)
             Possible.Impacts
nbr.val             683.000000
nbr.null              0.000000
nbr.na                0.000000
min                   1.000000
max                1144.000000
range              1143.000000
sum               29607.000000
median                6.000000
mean                 43.348463
SE.mean               4.353414
CI.mean.0.95          8.547703
var               12944.359334
std.dev             113.773280
coef.var              2.624621
             Cumulative.Impact.Probability
nbr.val                       6.830000e+02
nbr.null                      0.000000e+00
nbr.na                        0.000000e+00
min                           1.100000e-10
max                           6.500000e-02
range                         6.500000e-02
```

impacts

Filter

ity	Asteroid.Velocity	Asteroid.Magnitude	Asteroid.Diameter..km.	Cumulative.Palermo.Sc
	17.77	28.3	0.007	-8.31
	8.98	31.4	0.002	-6.60
	18.33	31.4	0.002	-6.48
	4.99	26.7	0.016	-6.83
	19.46	19.2	0.497	-3.85
	5.98	22.5	0.110	-5.03

Showing 1 to 8 of 683 entries, 11 total columns

Console | Terminal | Jobs

~/

```
var           2.178974e+00          NA
std.dev       1.476135e+00          NA
coef.var     -2.169507e-01          NA
> boxplot(pos.impacts)
> boxplot(pos.impacts, main= "Box Plot for Possible Impacts", ylab=
 "No. of Possible Impacts")
>
```

Environment | History | Connections

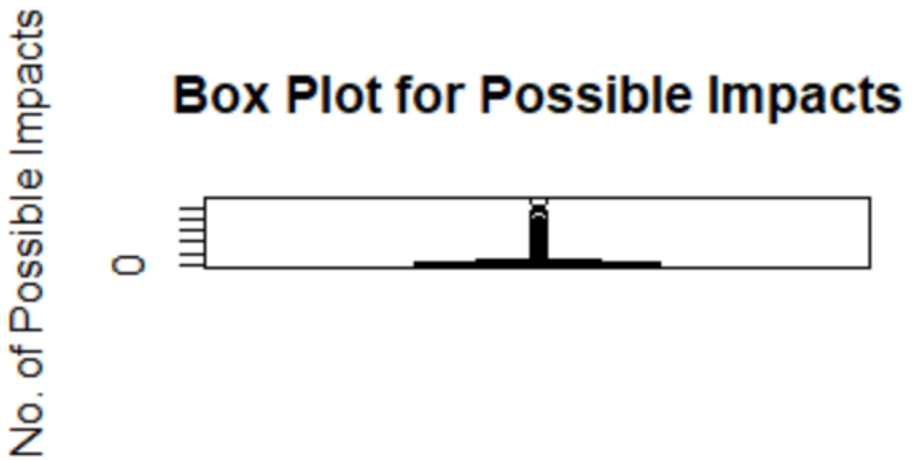
Import Dataset | List

Global Environment

cps	num [1:683] -8.31 -6.6 -6.4...
cum.impact...	num [1:683] 5.2e-09 7.6e-05...
cum.palermo	num [1:683] -8.31 -6.6 -6.4...
imp	int [1:683] 1 23 30 24 85 5...
keeps	int [1:683] 2 2 2 2 2 2 2 2...
max.palermo	num [1:683] -8.31 -6.96 -6....

Files | Plots | Packages | Help | Viewer

Zoom | Export



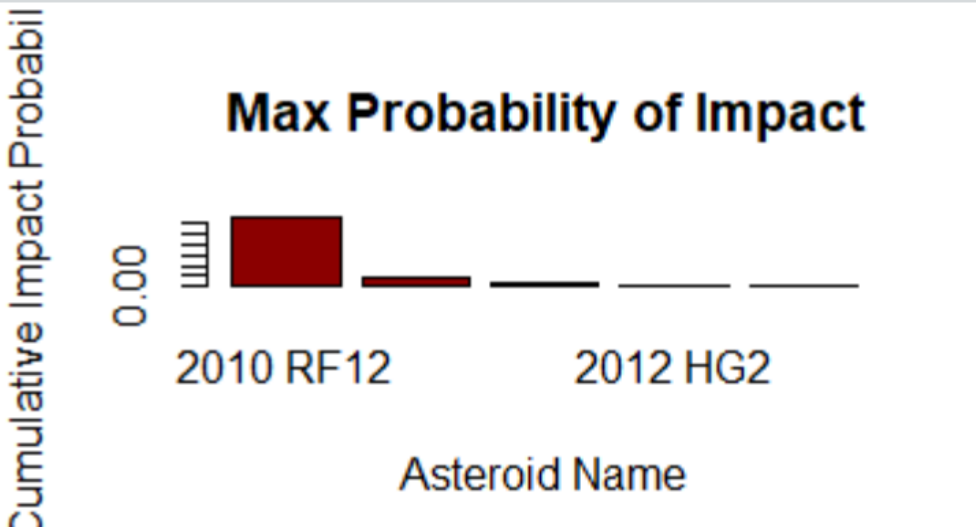
Console | Terminal | Jobs

```
> barplot(c(0.065, 0.0097, 0.0022, 0.002, 0.0019), main="Max Probab
ility of Impact", xlab= "Asteroid Name", ylab= "Cumulative Impact P
robability", names.arg= c("2010 RF12", "2006 JY26", "2000 SG334",
 "2012 HG2", "2016 RD34", col = "darkred"))
Error in barplot.default(c(0.065, 0.0097, 0.0022, 0.002, 0.0019), m
ain = "Max Probability of Impact",  :
  incorrect number of names
> barplot(c(0.065, 0.0097, 0.0022, 0.002, 0.0019), main="Max Probab
ility of Impact", xlab= "Asteroid Name", ylab= "Cumulative Impact P
robability", names.arg= c("2010 RF12", "2006 JY26", "2000 SG334",
 "2012 HG2", "2016 RD34"), col = "darkred"))
Error: unexpected ')' in "barplot(c(0.065, 0.0097, 0.0022, 0.002,
 0.0019), main="Max Probability of Impact", xlab= "Asteroid Name",
 ylab= "Cumulative Impact Probability", names.arg= c("2010 RF12",
 "2006 JY26", "2000 SG"
> barplot(c(0.065, 0.0097, 0.0022, 0.002, 0.0019), main="Max Probab
ility of Impact", xlab= "Asteroid Name", ylab= "Cumulative Impact P
robability", names.arg= c("2010 RF12", "2006 JY26", "2000 SG334",
 "2012 HG2", "2016 RD34"), col= "darkred")
>
```

mp	num [1:683] -8.31 -6.96 -6.…
obj.name	Factor w/ 683 levels "10195…
period.end	int [1:683] 2017 2046 2062 …
period.sta…	int [1:683] 2017 2017 2017 …
pos.impacts	int [1:683] 1 23 30 24 85 5…
top.prob	logi [1:683] FALSE FALSE FA…

Files | Plots | Packages | Help | Viewer

Zoom | Export




```
> max(cum.impact.prob)
[1] 0.065
> library(dplyr)
Error in library(dplyr) : there is no package called ‘dplyr’
> max(cum.impact.prob >= 0.0019)
[1] 1
> cum.impact.prob >= 0.0019
```

Source

Console | Terminal | Jobs

~/

```
> cor.test(asteroid.diam, pos.impacts,method=c("pearson"))

        Pearson's product-moment correlation

data:  asteroid.diam and pos.impacts
t = -0.88977, df = 681, p-value = 0.3739
alternative hypothesis: true correlation is not equal to 0
95 percent confidence interval:
 -0.10881793  0.04104877
sample estimates:
        cor
-0.03407614

> cor.test(cum.palermo, pos.impacts,method=c("pearson"))

        Pearson's product-moment correlation

data:  cum.palermo and pos.impacts
t = 4.4556, df = 681, p-value = 9.778e-06
alternative hypothesis: true correlation is not equal to 0
95 percent confidence interval:
 0.09447703 0.24029014
```

Source

Console | Terminal | Jobs

~/

```
> model = lm(pos.impacts ~ cum.impact.prob + asteroid.vel + asteroi
d.mag + asteroid.diam + cum.palermo + max.palermo)
> summary(model)

Call:
lm(formula = pos.impacts ~ cum.impact.prob + asteroid.vel + asteroi
d.mag +
    asteroid.diam + cum.palermo + max.palermo)

Residuals:
    Min      1Q  Median      3Q     Max
-150.55  -40.12   -1.01   20.88 1079.49

Coefficients:
                 Estimate Std. Error t value Pr(>|t|)
(Intercept)      -63.3638    54.8264  -1.156    0.248
cum.impact.prob  775.0668  1427.0326   0.543    0.587
asteroid.vel      -0.2493     0.6295  -0.396    0.692
asteroid.mag       3.2995     2.0653   1.598    0.111
asteroid.diam      3.0512    28.6440   0.107    0.915
cum.palermo      212.4975    11.6770  18.198   <2e-16 ***
max.palermo     -206.7428    12.0644 -17.137   <2e-16 ***
---
```

impacts

Filter

	Object.Name	Period.Start	Period.End	Possible.Impacts	Cumulative.Impac
174	2011 SR52	2034	2115	4	7.6e-10
59	2010 AU118	2020	2112	38	1.8e-08
683	29075 (1950 DA)	2880	2880	1	1.2e-04
226	2017 CH1	2044	2044	1	2.3e-10

Showing 1 to 5 of 683 entries, 11 total columns

Console | Terminal | Jobs

```
 "2005 GV190"),col=99CCCC)
Error: unexpected symbol in "barplot(c(2.579, 1.900, 1.300, 0.900,
 0.870), main= "Largest Asteroids", xlab="Asteroid Name", ylab= "As
teroid Diameter (km)", names.arg = c("2011 SR52", "2010 AU118", "29
075(1950 DA)", "2017 "
> barplot(c(2.579, 1.900, 1.300, 0.900, 0.870), main= "Largest Aste
roids", xlab="Asteroid Name", ylab= "Asteroid Diameter (km)", name
s.arg = c("2011 SR52", "2010 AU118", "29075(1950 DA)", "2017 CH1",
 "2005 GV190"),col=c("#009999"))
>
```

Environment | History | Connections

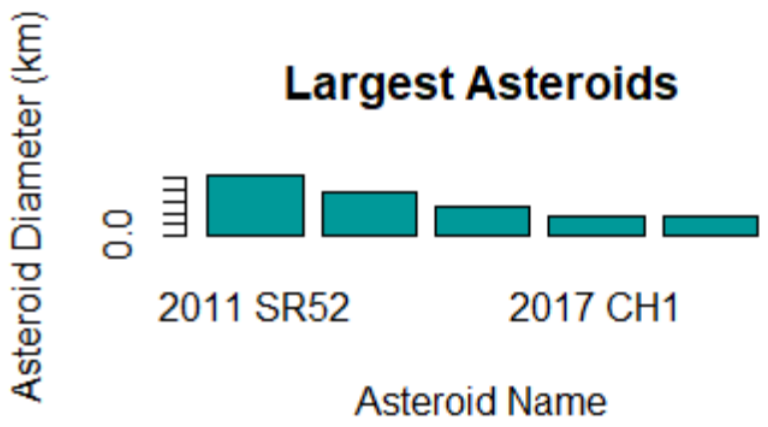
Import Dataset | List

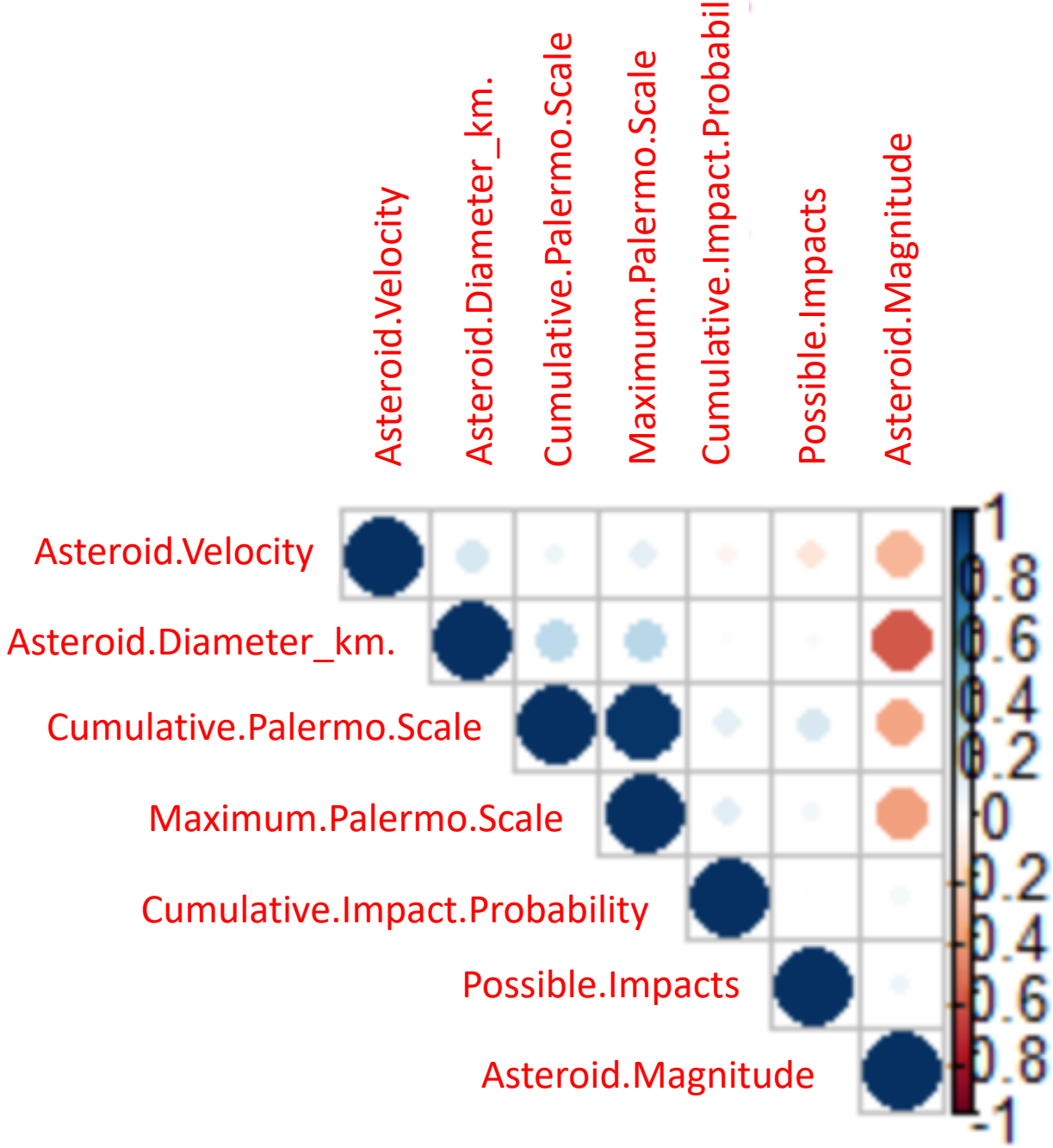
Global Environment

model	List of 12
Values	
ad	num [1:683] 0.007 0.002 0.0…
am	num [1:683] 28.3 31.4 31.4 …
asteroid.d…	num [1:683] 0.007 0.002 0.0…
asteroid.m…	num [1:683] 28.3 31.4 31.4 …

Files | Plots | Packages | Help | Viewer

Zoom | Export




Asteroid.Velocity
Asteroid.Diameter_km.
Cumulative.Palermo.Scale
Maximum.Palermo.Scale
Cumulative.Impact.Probabil
Possible.Impacts
Asteroid.Magnitude
Asteroid.Velocity
Asteroid.Diameter_km.
Cumulative.Palermo.Scale
Maximum.Palermo.Scale
Cumulative.Impact.Probability
Possible.Impacts
Asteroid.Magnitude
1
0.8
0.6
0.4
0.2
0
-0.2
-0.4
-0.6
-0.8
-1