

FSE_Survival

FSE Survival Analysis

This is a supplementary markdown file for FSE Submission title “Heard it through the GitVine: An Empirical Study of Tool Diffusion Across the npm Ecosystem.”

Defining Functions

```
get_df <- function(badge) {
  df <- read.csv(toString(badge), sep=",", stringsAsFactors = FALSE, header=FALSE)
  names(df) = c("repo_id", "start_interval", "end_interval", "act_committers", "act_pullers", "act_watchers",
               "tool_popularity", "all_popularity", "depsim", "depgraph", "desc", "n_commits", "n_commits_per_repo",
               "n_watchers", "n_org", "comp_badge", "other_badge", "adopted")

  df$adopted = as.factor(df$adopted)
  df$repo_id = as.factor(df$repo_id)
  df_stats = sqldf("select repo_id, count(*) as num_windows, max(adopted) as adopted from df group by repo_id")
  print(nrow(df_stats))

  summary(df_stats$num_windows)
  table(df_stats$num_windows)

  df_treatment = df_stats[df_stats$adopted == 1,]
  df_control = df_stats[df_stats$adopted == 0,][sample(nrow(df_stats), 2*nrow(df_treatment)),]

  print(nrow(df_control))
  print(nrow(df_treatment))

  df.sample = rbind(df[df$repo_id %in% df_treatment$repo_id,],
                   df[df$repo_id %in% df_control$repo_id,])

  print("Sample Created")
  print(nrow(df.sample))
  print(table(df.sample$adopted))
  summary(df.sample[df.sample$adopted==1,]$start_interval)
  return(df.sample)
}
```

Getting relevant data for badges.

```
df_travis <- get_df("travis")
df_circle <- get_df("circle")
df_appveyor <- get_df("appveyor")
df_codeship <- get_df("codeship")
df_david <- get_df("david")
df_bithound <- get_df("bithound")
df_gemnasium <- get_df("gemnasium")
```

```

df_coveralls <- get_df("coveralls")
df_codeclimate <- get_df("codeclimate")
df_codecov <- get_df("codecov")
df_codacy <- get_df("codacy")
df_saucelabs <- get_df("saucelabs")
df_slack <- get_df("slack")
df_gitter <- get_df("gitter")
df_license <- get_df("license")
df_npm_version <- get_df("npm-version")
df_npm_downloads <- get_df("npm-downloads")

```

Running Survival Models – David

```

coxph_david <- coxph(Surv(start_interval,
                          end_interval,
                          adopted == 1) ~
                    + log(n_commits+1)
                    + log(n_watchers+1)
                    + (other_badge > 0)
                    + log(all_popularity+1)
                    + (comp_badge>0)
                    + log(act_watchers+1)
                    + log(act_undir_committers+1)
                    + log(depgraph+1)
                    + log(depsim+1)
                    + log(desc+1)
                    , data=df_david)

summary(coxph_david)

```

```

## Call:
## coxph(formula = Surv(start_interval, end_interval, adopted ==
##      1) ~ +log(n_commits + 1) + log(n_watchers + 1) + (other_badge >
##      0) + log(all_popularity + 1) + (comp_badge > 0) + log(act_watchers +
##      1) + log(act_undir_committers + 1) + log(depgraph + 1) +
##      log(depsim + 1) + log(desc + 1), data = df_david)
##
##      n= 597016, number of events= 9798
##
##              coef exp(coef) se(coef)      z
## log(n_commits + 1)      0.346560  1.414195  0.009636  35.967
## log(n_watchers + 1)     -0.125004  0.882493  0.008953 -13.962
## other_badge > 0TRUE      0.972409  2.644306  0.025988  37.418
## log(all_popularity + 1)  0.015605  1.015727  0.011069   1.410
## comp_badge > 0TRUE     -2.894025  0.055353  0.707242  -4.092
## log(act_watchers + 1)   -0.031724  0.968774  0.008395  -3.779
## log(act_undir_committers + 1) 0.537605  1.711902  0.008500  63.248
## log(depgraph + 1)       0.753760  2.124974  0.024856  30.325
## log(depsim + 1)         0.130955  1.139916  0.039213   3.340
## log(desc + 1)          0.627245  1.872445  0.035165  17.837
##
##              Pr(>|z|)
## log(n_commits + 1)      < 2e-16 ***
## log(n_watchers + 1)     < 2e-16 ***

```

```
## other_badge > 0TRUE          < 2e-16 ***
## log(all_popularity + 1)      0.158608
## comp_badge > 0TRUE          4.28e-05 ***
## log(act_watchers + 1)        0.000157 ***
## log(act_undir_committers + 1) < 2e-16 ***
## log(depgraph + 1)           < 2e-16 ***
## log(depsim + 1)              0.000839 ***
## log(desc + 1)                < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
##
##               exp(coef) exp(-coef) lower .95 upper .95
## log(n_commits + 1)      1.41419    0.7071    1.38774    1.4412
## log(n_watchers + 1)     0.88249    1.1332    0.86714    0.8981
## other_badge > 0TRUE     2.64431    0.3782    2.51299    2.7825
## log(all_popularity + 1) 1.01573    0.9845    0.99393    1.0380
## comp_badge > 0TRUE      0.05535   18.0659    0.01384    0.2214
## log(act_watchers + 1)    0.96877    1.0322    0.95296    0.9848
## log(act_undir_committers + 1) 1.71190    0.5841    1.68362    1.7407
## log(depgraph + 1)       2.12497    0.4706    2.02393    2.2311
## log(depsim + 1)         1.13992    0.8773    1.05559    1.2310
## log(desc + 1)           1.87244    0.5341    1.74774    2.0060
##
## Concordance= 0.784 (se = 0.003 )
## Likelihood ratio test= 10339 on 10 df,  p=<2e-16
## Wald test              = 13231 on 10 df,  p=<2e-16
## Score (logrank) test = 15923 on 10 df,  p=<2e-16
#vif(coxph_david)
#anova(coxph_david)
```

Running Survival Models – Bithound

```
coxph_bithound <- coxph(Surv(start_interval,
                             end_interval,
                             adopted == 1) ~
  + log(n_commits+1)
  + log(n_watchers+1)
  + (other_badge > 0)
  + log(all_popularity+1)
  + (comp_badge>0)
  + log(act_watchers+1)
  + log(act_undir_committers+1)
  + log(depgraph+1)
  #+ log(depsim+1)
  + log(desc+1)
  , data=df_bithound)
summary(coxph_bithound)

## Call:
## coxph(formula = Surv(start_interval, end_interval, adopted ==
##      1) ~ +log(n_commits + 1) + log(n_watchers + 1) + (other_badge >
##      0) + log(all_popularity + 1) + (comp_badge > 0) + log(act_watchers +
```

```
##      1) + log(act_undir_committers + 1) + log(depgraph + 1) +
##      log(desc + 1), data = df_bithound)
##
##      n= 49420, number of events= 744
##
##
##              coef exp(coef) se(coef)      z Pr(>|z|)
## log(n_commits + 1)      0.45594   1.57766  0.03591 12.698 < 2e-16
## log(n_watchers + 1)     -0.02130   0.97892  0.02992 -0.712  0.477
## other_badge > OTRUE      0.79282   2.20962  0.10478  7.566 3.84e-14
## log(all_popularity + 1)  0.87474   2.39825  0.08293 10.548 < 2e-16
## comp_badge > OTRUE      0.23918   1.27020  0.44925  0.532  0.594
## log(act_watchers + 1)    -0.07508   0.92767  0.06618 -1.135  0.257
## log(act_undir_committers + 1) 1.05907   2.88368  0.05355 19.778 < 2e-16
## log(depgraph + 1)       0.06538   1.06757  0.18284  0.358  0.721
## log(desc + 1)          1.67946   5.36267  0.25752  6.522 6.95e-11
##
## log(n_commits + 1)      ***
## log(n_watchers + 1)
## other_badge > OTRUE      ***
## log(all_popularity + 1)  ***
## comp_badge > OTRUE
## log(act_watchers + 1)
## log(act_undir_committers + 1) ***
## log(depgraph + 1)
## log(desc + 1)          ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
##              exp(coef) exp(-coef) lower .95 upper .95
## log(n_commits + 1)      1.5777      0.6339   1.4704   1.693
## log(n_watchers + 1)     0.9789      1.0215   0.9232   1.038
## other_badge > OTRUE      2.2096      0.4526   1.7994   2.713
## log(all_popularity + 1)  2.3983      0.4170   2.0385   2.822
## comp_badge > OTRUE      1.2702      0.7873   0.5266   3.064
## log(act_watchers + 1)    0.9277      1.0780   0.8148   1.056
## log(act_undir_committers + 1) 2.8837      0.3468   2.5964   3.203
## log(depgraph + 1)       1.0676      0.9367   0.7460   1.528
## log(desc + 1)          5.3627      0.1865   3.2373   8.883
##
## Concordance= 0.816 (se = 0.009 )
## Likelihood ratio test= 1171 on 9 df, p=<2e-16
## Wald test              = 1481 on 9 df, p=<2e-16
## Score (logrank) test = 3120 on 9 df, p=<2e-16
##
## vif(coxph_bithound)
## anova(coxph_bithound)
```

Running Survival Models – Gemnasium

```
coxph_gemnasium <- coxph(Surv(start_interval,
                             end_interval,
                             adopted == 1) ~
  + log(n_commits+1)
```

```

+ log(n_watchers+1)
+ (other_badge > 0)
+ log(all_popularity+1)
#+ (comp_badge>0)
+ log(act_watchers+1)
+ log(act_undir_committers+1)
+ log(depgraph+1)
+ log(depsim+1)
+ log(desc+1)
, data=df_gemnasium)
summary(coxph_gemnasium)

## Call:
## coxph(formula = Surv(start_interval, end_interval, adopted ==
##      1) ~ +log(n_commits + 1) + log(n_watchers + 1) + (other_badge >
##      0) + log(all_popularity + 1) + log(act_watchers + 1) + log(act_undir_committers +
##      1) + log(depgraph + 1) + log(depsim + 1) + log(desc + 1),
##      data = df_gemnasium)
##
##      n= 80039, number of events= 1240
##
##              coef exp(coef) se(coef)      z Pr(>|z|)
## log(n_commits + 1)      0.33151  1.39307  0.02667  12.429 < 2e-16
## log(n_watchers + 1)     -0.19847  0.81998  0.02694  -7.367 1.74e-13
## other_badge > 0TRUE      1.16519  3.20652  0.07600  15.332 < 2e-16
## log(all_popularity + 1)  -0.25955  0.77140  0.02051 -12.652 < 2e-16
## log(act_watchers + 1)    -0.01243  0.98765  0.03551  -0.350  0.726
## log(act_undir_committers + 1) 0.82207  2.27521  0.02649  31.028 < 2e-16
## log(depgraph + 1)        0.88273  2.41749  0.09666   9.132 < 2e-16
## log(depsim + 1)          0.31062  1.36428  0.34274   0.906  0.365
## log(desc + 1)            0.84034  2.31714  0.13820   6.081 1.20e-09
##
## log(n_commits + 1)      ***
## log(n_watchers + 1)     ***
## other_badge > 0TRUE      ***
## log(all_popularity + 1)  ***
## log(act_watchers + 1)
## log(act_undir_committers + 1) ***
## log(depgraph + 1)       ***
## log(depsim + 1)
## log(desc + 1)           ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
##              exp(coef) exp(-coef) lower .95 upper .95
## log(n_commits + 1)      1.3931  0.7178  1.3221  1.4678
## log(n_watchers + 1)      0.8200  1.2195  0.7778  0.8644
## other_badge > 0TRUE      3.2065  0.3119  2.7628  3.7215
## log(all_popularity + 1)  0.7714  1.2963  0.7410  0.8030
## log(act_watchers + 1)    0.9876  1.0125  0.9212  1.0588
## log(act_undir_committers + 1) 2.2752  0.4395  2.1601  2.3965
## log(depgraph + 1)        2.4175  0.4137  2.0003  2.9217
## log(depsim + 1)          1.3643  0.7330  0.6969  2.6708
## log(desc + 1)            2.3171  0.4316  1.7673  3.0380

```

```
##
## Concordance= 0.801 (se = 0.008 )
## Likelihood ratio test= 1698 on 9 df, p=<2e-16
## Wald test = 2487 on 9 df, p=<2e-16
## Score (logrank) test = 3926 on 9 df, p=<2e-16

#vif(coxph_gemnasium)
#anova(coxph_gemnasium)
```

Running Survival Models – Saucelabs

```
coxph_saucelabs <- coxph(Surv(start_interval,
                             end_interval,
                             adopted == 1) ~
+ log(n_commits+1)
+ log(n_watchers+1)
+ (other_badge > 0)
+ log(all_popularity+1)
+ (comp_badge>0)
+ log(act_watchers+1)
+ log(act_undir_committers+1)
+ log(depgraph+1)
+ log(depsim+1)
+ log(desc+1)
, data=df_saucelabs)
summary(coxph_saucelabs)
```

```
## Call:
## coxph(formula = Surv(start_interval, end_interval, adopted ==
## 1) ~ +log(n_commits + 1) + log(n_watchers + 1) + (other_badge >
## 0) + log(all_popularity + 1) + (comp_badge > 0) + log(act_watchers +
## 1) + log(act_undir_committers + 1) + log(depgraph + 1) +
## log(depsim + 1) + log(desc + 1), data = df_saucelabs)
##
## n= 37830, number of events= 541
##
##
```

	coef	exp(coef)	se(coef)	z	Pr(> z)
## log(n_commits + 1)	0.51712	1.67720	0.03825	13.519	< 2e-16
## log(n_watchers + 1)	-0.06673	0.93545	0.02860	-2.333	0.01963
## other_badge > 0TRUE	1.12374	3.07634	0.12263	9.164	< 2e-16
## log(all_popularity + 1)	-0.10798	0.89764	0.04036	-2.675	0.00746
## comp_badge > 0TRUE	-1.12205	0.32561	1.00144	-1.120	0.26253
## log(act_watchers + 1)	0.10987	1.11614	0.05130	2.142	0.03222
## log(act_undir_committers + 1)	0.78955	2.20240	0.07149	11.044	< 2e-16
## log(depgraph + 1)	1.30485	3.68714	0.16132	8.089	6.02e-16
## log(depsim + 1)	0.24484	1.27742	1.05448	0.232	0.81639
## log(desc + 1)	2.46189	11.72699	0.53083	4.638	3.52e-06

```
##
## log(n_commits + 1) ***
## log(n_watchers + 1) *
## other_badge > 0TRUE ***
## log(all_popularity + 1) **
## comp_badge > 0TRUE
```

```
## log(act_watchers + 1) *
## log(act_undir_committers + 1) ***
## log(depgraph + 1) ***
## log(depsim + 1)
## log(desc + 1) ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
##               exp(coef) exp(-coef) lower .95 upper .95
## log(n_commits + 1)      1.6772    0.59623    1.55605    1.8078
## log(n_watchers + 1)     0.9354    1.06901    0.88446    0.9894
## other_badge > 0TRUE     3.0763    0.32506    2.41908    3.9122
## log(all_popularity + 1)  0.8976    1.11403    0.82937    0.9715
## comp_badge > 0TRUE      0.3256    3.07114    0.04574    2.3181
## log(act_watchers + 1)   1.1161    0.89595    1.00937    1.2342
## log(act_undir_committers + 1) 2.2024    0.45405    1.91445    2.5337
## log(depgraph + 1)      3.6871    0.27121    2.68768    5.0583
## log(depsim + 1)        1.2774    0.78283    0.16172    10.0904
## log(desc + 1)          11.7270    0.08527    4.14326    33.1918
##
## Concordance= 0.785  (se = 0.014 )
## Likelihood ratio test= 654.9  on 10 df,  p=<2e-16
## Wald test              = 808.5  on 10 df,  p=<2e-16
## Score (logrank) test = 1016   on 10 df,  p=<2e-16
#vif(coxph_saucelabs)
#anova(coxph_saucelabs)
```

Running Survival Models – slack

```
coxph_slack <- coxph(Surv(start_interval,
                          end_interval,
                          adopted == 1) ~
  + log(n_commits+1)
  + log(n_watchers+1)
  + (other_badge > 0)
  + log(all_popularity+1)
  #+ (comp_badge>0)
  + log(act_watchers+1)
  + log(act_undir_committers+1)
  + log(depgraph+1)
  + log(depsim+1)
  + log(desc+1)
  , data=df_slack)
summary(coxph_slack)

## Call:
## coxph(formula = Surv(start_interval, end_interval, adopted ==
##      1) ~ +log(n_commits + 1) + log(n_watchers + 1) + (other_badge >
##      0) + log(all_popularity + 1) + log(act_watchers + 1) + log(act_undir_committers +
##      1) + log(depgraph + 1) + log(depsim + 1) + log(desc + 1),
##      data = df_slack)
##
```

```
## n= 35001, number of events= 445
##
##
##          coef exp(coef) se(coef)      z Pr(>|z|)
## log(n_commits + 1)      0.08075   1.08409  0.04363  1.851  0.0642
## log(n_watchers + 1)     0.14891   1.16057  0.03064  4.860 1.17e-06
## other_badge > OTRUE     0.86773   2.38151  0.16301  5.323 1.02e-07
## log(all_popularity + 1)  0.92854   2.53080  0.13018  7.132 9.86e-13
## log(act_watchers + 1)    0.09566   1.10039  0.05592  1.711  0.0871
## log(act_undir_committers + 1) 1.00849   2.74145  0.04765 21.165 < 2e-16
## log(depgraph + 1)       0.13491   1.14443  0.14294  0.944  0.3453
## log(depsim + 1)         0.04159   1.04246  0.15584  0.267  0.7896
## log(desc + 1)          0.07368   1.07646  0.08665  0.850  0.3952
##
## log(n_commits + 1)      .
## log(n_watchers + 1)     ***
## other_badge > OTRUE     ***
## log(all_popularity + 1) ***
## log(act_watchers + 1)    .
## log(act_undir_committers + 1) ***
## log(depgraph + 1)
## log(depsim + 1)
## log(desc + 1)
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
##          exp(coef) exp(-coef) lower .95 upper .95
## log(n_commits + 1)      1.084      0.9224   0.9952   1.181
## log(n_watchers + 1)     1.161      0.8616   1.0929   1.232
## other_badge > OTRUE     2.382      0.4199   1.7302   3.278
## log(all_popularity + 1)  2.531      0.3951   1.9609   3.266
## log(act_watchers + 1)    1.100      0.9088   0.9862   1.228
## log(act_undir_committers + 1) 2.741      0.3648   2.4970   3.010
## log(depgraph + 1)       1.144      0.8738   0.8648   1.514
## log(depsim + 1)         1.042      0.9593   0.7681   1.415
## log(desc + 1)          1.076      0.9290   0.9083   1.276
##
## Concordance= 0.889 (se = 0.01 )
## Likelihood ratio test= 1130 on 9 df, p=<2e-16
## Wald test              = 1454 on 9 df, p=<2e-16
## Score (logrank) test = 4656 on 9 df, p=<2e-16
##
##vif(coxph_slack)
##anova(coxph_slack)
```

Running Survival Models – gitter

```
coxph_gitter <- coxph(Surv(start_interval,
                           end_interval,
                           adopted == 1) ~
  + log(n_commits+1)
  + log(n_watchers+1)
  + (other_badge > 0)
  + log(all_popularity+1))
```



```

      #+ (comp_badge>0)
      + log(act_watchers+1)
      + log(act_undir_committers+1)
      + log(depgraph+1)
      + log(depsim+1)
      + log(desc+1)
      , data=df_gitter)
summary(coxph_gitter)

## Call:
## coxph(formula = Surv(start_interval, end_interval, adopted ==
##      1) ~ +log(n_commits + 1) + log(n_watchers + 1) + (other_badge >
##      0) + log(all_popularity + 1) + log(act_watchers + 1) + log(act_undir_committers +
##      1) + log(depgraph + 1) + log(depsim + 1) + log(desc + 1),
##      data = df_gitter)
##
##      n= 204006, number of events= 2957
##
##               coef exp(coef) se(coef)      z Pr(>|z|)
## log(n_commits + 1)      0.32839   1.38873  0.01578 20.809 < 2e-16
## log(n_watchers + 1)      0.14106   1.15149  0.01167 12.090 < 2e-16
## other_badge > OTRUE      0.45485   1.57593  0.04658  9.765 < 2e-16
## log(all_popularity + 1)  0.29075   1.33743  0.02738 10.620 < 2e-16
## log(act_watchers + 1)   -0.04427   0.95670  0.01527 -2.899 0.00374
## log(act_undir_committers + 1) 0.57639   1.77960  0.01973 29.215 < 2e-16
## log(depgraph + 1)       0.97215   2.64361  0.05289 18.381 < 2e-16
## log(depsim + 1)        -0.15835   0.85355  0.13311 -1.190 0.23421
## log(desc + 1)          1.05053   2.85917  0.07224 14.542 < 2e-16
##
## log(n_commits + 1)      ***
## log(n_watchers + 1)      ***
## other_badge > OTRUE      ***
## log(all_popularity + 1)  ***
## log(act_watchers + 1)    **
## log(act_undir_committers + 1) ***
## log(depgraph + 1)       ***
## log(depsim + 1)         ***
## log(desc + 1)           ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
##               exp(coef) exp(-coef) lower .95 upper .95
## log(n_commits + 1)      1.3887      0.7201      1.3464      1.4324
## log(n_watchers + 1)      1.1515      0.8684      1.1255      1.1781
## other_badge > OTRUE      1.5759      0.6345      1.4384      1.7266
## log(all_popularity + 1)  1.3374      0.7477      1.2676      1.4112
## log(act_watchers + 1)    0.9567      1.0453      0.9285      0.9858
## log(act_undir_committers + 1) 1.7796      0.5619      1.7121      1.8498
## log(depgraph + 1)       2.6436      0.3783      2.3833      2.9324
## log(depsim + 1)         0.8536      1.1716      0.6575      1.1080
## log(desc + 1)           2.8592      0.3498      2.4817      3.2941
##
## Concordance= 0.789 (se = 0.005 )
## Likelihood ratio test= 3543 on 9 df,  p=<2e-16

```

```
## Wald test          = 5185  on 9 df,   p=<2e-16
## Score (logrank) test = 7934  on 9 df,   p=<2e-16
```

```
#vif(coxph_gitter)
#anova(coxph_gitter)
```

Running Survival Models – license

```
coxph_license <- coxph(Surv(start_interval,
                             end_interval,
                             adopted == 1) ~
  + log(n_commits+1)
  + log(n_watchers+1)
  + (other_badge > 0)
  + log(all_popularity+1)
  #+ (comp_badge>0)
  + log(act_watchers+1)
  + log(act_undir_committers+1)
  + log(depgraph+1)
  + log(depsim+1)
  + log(desc+1)
  , data=df_license)

summary(coxph_license)
```

```
## Call:
## coxph(formula = Surv(start_interval, end_interval, adopted ==
##      1) ~ +log(n_commits + 1) + log(n_watchers + 1) + (other_badge >
##      0) + log(all_popularity + 1) + log(act_watchers + 1) + log(act_undir_committers +
##      1) + log(depgraph + 1) + log(depsim + 1) + log(desc + 1),
##      data = df_license)
##
##      n= 189526, number of events= 2886
##
##              coef exp(coef) se(coef)      z Pr(>|z|)
## log(n_commits + 1)      0.40904   1.50537  0.01733 23.609 <2e-16
## log(n_watchers + 1)     -0.03315   0.96739  0.01530 -2.167  0.0302
## other_badge > 0TRUE      0.78673   2.19620  0.04831 16.283 <2e-16
## log(all_popularity + 1)  0.85224   2.34489  0.03856 22.102 <2e-16
## log(act_watchers + 1)    -0.21023   0.81040  0.02297 -9.154 <2e-16
## log(act_undir_committers + 1) 0.87664   2.40280  0.02164 40.509 <2e-16
## log(depgraph + 1)       0.70870   2.03135  0.06934 10.221 <2e-16
## log(depsim + 1)         0.22674   1.25450  0.12229  1.854  0.0637
## log(desc + 1)          0.72979   2.07465  0.06230 11.714 <2e-16
##
## log(n_commits + 1)      ***
## log(n_watchers + 1)      *
## other_badge > 0TRUE      ***
## log(all_popularity + 1)  ***
## log(act_watchers + 1)    ***
## log(act_undir_committers + 1) ***
## log(depgraph + 1)       ***
## log(depsim + 1)         .
## log(desc + 1)          ***
```

```
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
##
##               exp(coef) exp(-coef) lower .95 upper .95
## log(n_commits + 1)      1.5054      0.6643      1.4551      1.5574
## log(n_watchers + 1)     0.9674      1.0337      0.9388      0.9968
## other_badge > 0TRUE     2.1962      0.4553      1.9978      2.4143
## log(all_popularity + 1)  2.3449      0.4265      2.1742      2.5290
## log(act_watchers + 1)    0.8104      1.2340      0.7747      0.8477
## log(act_undir_committers + 1) 2.4028      0.4162      2.3030      2.5069
## log(depgraph + 1)       2.0314      0.4923      1.7732      2.3271
## log(depsim + 1)         1.2545      0.7971      0.9871      1.5943
## log(desc + 1)          2.0746      0.4820      1.8362      2.3441
##
## Concordance= 0.824 (se = 0.005 )
## Likelihood ratio test= 4745 on 9 df,  p=<2e-16
## Wald test              = 6324 on 9 df,  p=<2e-16
## Score (logrank) test = 11646 on 9 df,  p=<2e-16
vif(coxph_license)

## Warning in vif.default(coxph_license): No intercept: vifs may not be
## sensible.

##           log(n_commits + 1)           log(n_watchers + 1)
##           1.228838           1.191832
##           other_badge > 0           log(all_popularity + 1)
##           1.069011           1.040604
##           log(act_watchers + 1) log(act_undir_committers + 1)
##           1.420600           1.786004
##           log(depgraph + 1)           log(depsim + 1)
##           1.272154           1.006543
##           log(desc + 1)
##           1.125196

anova(coxph_license)

## Analysis of Deviance Table
## Cox model: response is Surv(start_interval, end_interval, adopted == 1)
## Terms added sequentially (first to last)
##
##               loglik      Chisq Df Pr(>|Chi|)
## NULL
## -24860
## log(n_commits + 1) -24184 1350.3506  1 < 2.2e-16 ***
## log(n_watchers + 1) -24167  34.5969  1 4.056e-09 ***
## other_badge > 0 -23894 545.9824  1 < 2.2e-16 ***
## log(all_popularity + 1) -23489 809.9686  1 < 2.2e-16 ***
## log(act_watchers + 1) -23478  22.0044  1 2.720e-06 ***
## log(act_undir_committers + 1) -22584 1787.8387  1 < 2.2e-16 ***
## log(depgraph + 1) -22537  93.9066  1 < 2.2e-16 ***
## log(depsim + 1) -22536  3.1729  1 0.07487 .
## log(desc + 1) -22487 96.9463  1 < 2.2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Running Survival Models – npm-version

```
coxph_npm_version <- coxph(Surv(start_interval,
                               end_interval,
                               adopted == 1) ~
    + log(n_commits+1)
    + log(n_watchers+1)
    + (other_badge > 0)
    + log(all_popularity+1)
    #+ (comp_badge>0)
    + log(act_watchers+1)
    + log(act_undir_committers+1)
    + log(depgraph+1)
    + log(depsim+1)
    + log(desc+1)
    , data=df_npm_version)
summary(coxph_npm_version)
```

```
## Call:
## coxph(formula = Surv(start_interval, end_interval, adopted ==
##      1) ~ +log(n_commits + 1) + log(n_watchers + 1) + (other_badge >
##      0) + log(all_popularity + 1) + log(act_watchers + 1) + log(act_undir_committers +
##      1) + log(depgraph + 1) + log(depsim + 1) + log(desc + 1),
##      data = df_npm_version)
##
##      n= 1199367, number of events= 21985
##
##              coef exp(coef) se(coef)      z
## log(n_commits + 1)      0.3732243  1.4524101  0.0061749  60.443
## log(n_watchers + 1)     -0.0507032  0.9505607  0.0054816  -9.250
## other_badge > 0TRUE      0.8310680  2.2957692  0.0160511  51.777
## log(all_popularity + 1)  0.3062121  1.3582703  0.0095038  32.220
## log(act_watchers + 1)    -0.0191236  0.9810581  0.0047608  -4.017
## log(act_undir_committers + 1) 0.3968712  1.4871644  0.0056231  70.579
## log(depgraph + 1)       0.3818441  1.4649836  0.0153034  24.952
## log(depsim + 1)         0.0009155  1.0009159  0.0243387   0.038
## log(desc + 1)          0.4078782  1.5036239  0.0211932  19.246
##
##              Pr(>|z|)
## log(n_commits + 1)      < 2e-16 ***
## log(n_watchers + 1)      < 2e-16 ***
## other_badge > 0TRUE      < 2e-16 ***
## log(all_popularity + 1)  < 2e-16 ***
## log(act_watchers + 1)    5.9e-05 ***
## log(act_undir_committers + 1) < 2e-16 ***
## log(depgraph + 1)       < 2e-16 ***
## log(depsim + 1)         0.97
## log(desc + 1)          < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
##              exp(coef) exp(-coef) lower .95 upper .95
## log(n_commits + 1)      1.4524      0.6885      1.4349      1.4701
## log(n_watchers + 1)      0.9506      1.0520      0.9404      0.9608
## other_badge > 0TRUE      2.2958      0.4356      2.2247      2.3691
```

```
## log(all_popularity + 1)          1.3583      0.7362      1.3332      1.3838
## log(act_watchers + 1)           0.9811      1.0193      0.9719      0.9903
## log(act_undir_committers + 1)    1.4872      0.6724      1.4709      1.5036
## log(depgraph + 1)               1.4650      0.6826      1.4217      1.5096
## log(depsim + 1)                 1.0009      0.9991      0.9543      1.0498
## log(desc + 1)                   1.5036      0.6651      1.4424      1.5674
##
## Concordance= 0.783 (se = 0.002 )
## Likelihood ratio test= 21117 on 9 df,  p=<2e-16
## Wald test               = 23468 on 9 df,  p=<2e-16
## Score (logrank) test = 26384 on 9 df,  p=<2e-16

#vif(coxph_npm_version)
#anova(coxph_npm_version)
```

Running Survival Models – npm-downloads

```
coxph_npm_downloads <- coxph(Surv(start_interval,
                                   end_interval,
                                   adopted == 1) ~
  + log(n_commits+1)
  + log(n_watchers+1)
  + (other_badge > 0)
  + log(all_popularity+1)
  #+ (comp_badge>0)
  + log(act_watchers+1)
  + log(act_undir_committers+1)
  + log(depgraph+1)
  + log(depsim+1)
  + log(desc+1)
  , data=df_npm_downloads)
summary(coxph_npm_downloads)
```

```
## Call:
## coxph(formula = Surv(start_interval, end_interval, adopted ==
##      1) ~ +log(n_commits + 1) + log(n_watchers + 1) + (other_badge >
##      0) + log(all_popularity + 1) + log(act_watchers + 1) + log(act_undir_committers +
##      1) + log(depgraph + 1) + log(depsim + 1) + log(desc + 1),
##      data = df_npm_downloads)
##
##      n= 493629, number of events= 7594
##
##              coef exp(coef) se(coef)      z
## log(n_commits + 1)      0.268953  1.308593  0.010692 25.154
## log(n_watchers + 1)     0.050011  1.051283  0.008524  5.867
## other_badge > 0TRUE     0.916465  2.500435  0.031366 29.218
## log(all_popularity + 1)  0.500151  1.648970  0.020287 24.654
## log(act_watchers + 1)   -0.052044  0.949287  0.009378 -5.550
## log(act_undir_committers + 1) 0.582517  1.790539  0.009508 61.267
## log(depgraph + 1)       0.635859  1.888643  0.028849 22.041
## log(depsim + 1)        -0.015431  0.984687  0.078630 -0.196
## log(desc + 1)          0.835316  2.305542  0.037863 22.061
##
## Pr(>|z|)
```

```
## log(n_commits + 1) < 2e-16 ***
## log(n_watchers + 1) 4.44e-09 ***
## other_badge > 0TRUE < 2e-16 ***
## log(all_popularity + 1) < 2e-16 ***
## log(act_watchers + 1) 2.86e-08 ***
## log(act_undir_committers + 1) < 2e-16 ***
## log(depgraph + 1) < 2e-16 ***
## log(depsim + 1) 0.844
## log(desc + 1) < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
##
## exp(coef) exp(-coef) lower .95 upper .95
## log(n_commits + 1) 1.3086 0.7642 1.2815 1.3363
## log(n_watchers + 1) 1.0513 0.9512 1.0339 1.0690
## other_badge > 0TRUE 2.5004 0.3999 2.3513 2.6590
## log(all_popularity + 1) 1.6490 0.6064 1.5847 1.7159
## log(act_watchers + 1) 0.9493 1.0534 0.9320 0.9669
## log(act_undir_committers + 1) 1.7905 0.5585 1.7575 1.8242
## log(depgraph + 1) 1.8886 0.5295 1.7848 1.9985
## log(depsim + 1) 0.9847 1.0156 0.8441 1.1488
## log(desc + 1) 2.3055 0.4337 2.1406 2.4831
##
## Concordance= 0.803 (se = 0.003 )
## Likelihood ratio test= 9856 on 9 df, p=<2e-16
## Wald test = 13130 on 9 df, p=<2e-16
## Score (logrank) test = 18309 on 9 df, p=<2e-16
```

Running Survival Models – travis

```
coxph_travis <- coxph(Surv(start_interval,
                           end_interval,
                           adopted == 1) ~
  + log(n_commits+1)
  + log(n_watchers+1)
  + (other_badge > 0)
  + log(all_popularity+1)
  + (comp_badge>0)
  + log(act_watchers+1)
  + log(act_undir_committers+1)
  + log(depgraph+1)
  + log(depsim+1)
  + log(desc+1)
  , data=df_travis)
summary(coxph_travis)

## Call:
## coxph(formula = Surv(start_interval, end_interval, adopted ==
## 1) ~ +log(n_commits + 1) + log(n_watchers + 1) + (other_badge >
## 0) + log(all_popularity + 1) + (comp_badge > 0) + log(act_watchers +
## 1) + log(act_undir_committers + 1) + log(depgraph + 1) +
## log(depsim + 1) + log(desc + 1), data = df_travis)
##
```

```
## n= 1253861, number of events= 32283
##
##
##          coef exp(coef) se(coef)      z
## log(n_commits + 1)      0.458145  1.581138  0.004957  92.422
## log(n_watchers + 1)     -0.115533  0.890891  0.004613 -25.046
## other_badge > 0TRUE      0.673177  1.960456  0.012199  55.182
## log(all_popularity + 1)  -0.083987  0.919443  0.003862 -21.745
## comp_badge > 0TRUE      -4.423679  0.011990  0.316318 -13.985
## log(act_watchers + 1)    0.005011  1.005023  0.003645   1.375
## log(act_undir_committers + 1) 0.314742  1.369906  0.004482  70.224
## log(depgraph + 1)       0.160286  1.173847  0.009860  16.256
## log(depsim + 1)         -0.037299  0.963388  0.020701  -1.802
## log(desc + 1)           0.364683  1.440058  0.017110  21.315
##
##          Pr(>|z|)
## log(n_commits + 1)      <2e-16 ***
## log(n_watchers + 1)     <2e-16 ***
## other_badge > 0TRUE     <2e-16 ***
## log(all_popularity + 1) <2e-16 ***
## comp_badge > 0TRUE     <2e-16 ***
## log(act_watchers + 1)    0.1692
## log(act_undir_committers + 1) <2e-16 ***
## log(depgraph + 1)       <2e-16 ***
## log(depsim + 1)         0.0716 .
## log(desc + 1)           <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
##
##          exp(coef) exp(-coef) lower .95 upper .95
## log(n_commits + 1)      1.58114      0.6325      1.56585      1.59657
## log(n_watchers + 1)     0.89089      1.1225      0.88287      0.89898
## other_badge > 0TRUE     1.96046      0.5101      1.91414      2.00789
## log(all_popularity + 1)  0.91944      1.0876      0.91251      0.92643
## comp_badge > 0TRUE      0.01199     83.4026      0.00645      0.02229
## log(act_watchers + 1)    1.00502      0.9950      0.99787      1.01223
## log(act_undir_committers + 1) 1.36991      0.7300      1.35792      1.38199
## log(depgraph + 1)       1.17385      0.8519      1.15138      1.19675
## log(depsim + 1)         0.96339      1.0380      0.92508      1.00328
## log(desc + 1)           1.44006      0.6944      1.39257      1.48917
##
## Concordance= 0.77 (se = 0.002 )
## Likelihood ratio test= 24345 on 10 df, p=<2e-16
## Wald test              = 26671 on 10 df, p=<2e-16
## Score (logrank) test = 28088 on 10 df, p=<2e-16
```

Running Survival Models – circle

```
coxph_circle <- coxph(Surv(start_interval,
                           end_interval,
                           adopted == 1) ~
  + log(n_commits+1)
  + log(n_watchers+1)
  + (other_badge > 0)
```

```

+ log(all_popularity+1)
+ (comp_badge>0)
+ log(act_watchers+1)
+ log(act_undir_committers+1)
+ log(depgraph+1)
+ log(depsim+1)
+ log(desc+1)
, data=df_circle)
summary(coxph_circle)

```

```

## Call:
## coxph(formula = Surv(start_interval, end_interval, adopted ==
##      1) ~ +log(n_commits + 1) + log(n_watchers + 1) + (other_badge >
##      0) + log(all_popularity + 1) + (comp_badge > 0) + log(act_watchers +
##      1) + log(act_undir_committers + 1) + log(depgraph + 1) +
##      log(depsim + 1) + log(desc + 1), data = df_circle)
##
##      n= 121930, number of events= 1798
##
##              coef exp(coef) se(coef)      z Pr(>|z|)
## log(n_commits + 1)      0.37503   1.45504  0.02082  18.010 < 2e-16
## log(n_watchers + 1)    -0.07862   0.92439  0.01783  -4.410 1.03e-05
## other_badge > OTRUE      0.61942   1.85785  0.05204  11.904 < 2e-16
## log(all_popularity + 1)  0.83366   2.30173  0.05323  15.661 < 2e-16
## comp_badge > OTRUE     -2.47238   0.08438  0.12064 -20.494 < 2e-16
## log(act_watchers + 1)    0.06859   1.07100  0.02945   2.329  0.0199
## log(act_undir_committers + 1) 0.77986   2.18116  0.03337  23.372 < 2e-16
## log(depgraph + 1)       0.47539   1.60865  0.08448   5.627 1.83e-08
## log(depsim + 1)        -0.64351   0.52545  0.39633  -1.624  0.1044
## log(desc + 1)          0.59605   1.81494  0.11475   5.195 2.05e-07
##
## log(n_commits + 1)      ***
## log(n_watchers + 1)      ***
## other_badge > OTRUE      ***
## log(all_popularity + 1)  ***
## comp_badge > OTRUE      ***
## log(act_watchers + 1)    *
## log(act_undir_committers + 1) ***
## log(depgraph + 1)       ***
## log(depsim + 1)
## log(desc + 1)          ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
##              exp(coef) exp(-coef) lower .95 upper .95
## log(n_commits + 1)      1.45504   0.6873  1.39685  1.5157
## log(n_watchers + 1)    0.92439   1.0818  0.89265  0.9573
## other_badge > OTRUE      1.85785   0.5383  1.67771  2.0573
## log(all_popularity + 1)  2.30173   0.4345  2.07369  2.5548
## comp_badge > OTRUE      0.08438  11.8507  0.06661  0.1069
## log(act_watchers + 1)    1.07100   0.9337  1.01093  1.1346
## log(act_undir_committers + 1) 2.18116   0.4585  2.04308  2.3286
## log(depgraph + 1)       1.60865   0.6216  1.36317  1.8983
## log(depsim + 1)         0.52545   1.9031  0.24164  1.1426

```



```
## log(desc + 1)                1.81494      0.5510    1.44941    2.2727
##
## Concordance= 0.861 (se = 0.004 )
## Likelihood ratio test= 3371 on 10 df,    p=<2e-16
## Wald test              = 3637 on 10 df,    p=<2e-16
## Score (logrank) test = 6938 on 10 df,    p=<2e-16
```

Running Survival Models – appveyor

```
coxph_appveyor <- coxph(Surv(start_interval,
                             end_interval,
                             adopted == 1) ~
  + log(n_commits+1)
  + log(n_watchers+1)
  + (other_badge > 0)
  + log(all_popularity+1)
  + (comp_badge>0)
  + log(act_watchers+1)
  + log(act_undir_committers+1)
  + log(depgraph+1)
  + log(depsim+1)
  + log(desc+1)
  , data=df_appveyor)
summary(coxph_appveyor)
```

```
## Call:
## coxph(formula = Surv(start_interval, end_interval, adopted ==
##      1) ~ +log(n_commits + 1) + log(n_watchers + 1) + (other_badge >
##      0) + log(all_popularity + 1) + (comp_badge > 0) + log(act_watchers +
##      1) + log(act_undir_committers + 1) + log(depgraph + 1) +
##      log(depsim + 1) + log(desc + 1), data = df_appveyor)
##
##      n= 110958, number of events= 1598
##
##              coef exp(coef) se(coef)      z Pr(>|z|)
## log(n_commits + 1)      0.35726   1.42941  0.02299 15.538 < 2e-16
## log(n_watchers + 1)     -0.07999   0.92313  0.01763 -4.536 5.72e-06
## other_badge > OTRUE      1.51383   4.54411  0.07538 20.082 < 2e-16
## log(all_popularity + 1)  0.24086   1.27235  0.04137  5.822 5.83e-09
## comp_badge > OTRUE       0.02964   1.03009  0.05238  0.566  0.571
## log(act_watchers + 1)    0.11309   1.11973  0.02436  4.642 3.44e-06
## log(act_undir_committers + 1) 0.61172   1.84360  0.02647 23.111 < 2e-16
## log(depgraph + 1)        0.75942   2.13704  0.06944 10.936 < 2e-16
## log(depsim + 1)          0.90626   2.47504  0.22897  3.958 7.56e-05
## log(desc + 1)           1.48185   4.40110  0.14167 10.460 < 2e-16
##
## log(n_commits + 1)      ***
## log(n_watchers + 1)     ***
## other_badge > OTRUE      ***
## log(all_popularity + 1)  ***
## comp_badge > OTRUE
## log(act_watchers + 1)    ***
## log(act_undir_committers + 1) ***
```

```
## log(depgraph + 1)          ***
## log(depsim + 1)           ***
## log(desc + 1)             ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
##               exp(coef) exp(-coef) lower .95 upper .95
## log(n_commits + 1)      1.4294    0.6996    1.3664    1.4953
## log(n_watchers + 1)     0.9231    1.0833    0.8918    0.9556
## other_badge > OTRUE     4.5441    0.2201    3.9200    5.2676
## log(all_popularity + 1) 1.2723    0.7859    1.1732    1.3798
## comp_badge > OTRUE      1.0301    0.9708    0.9296    1.1415
## log(act_watchers + 1)   1.1197    0.8931    1.0675    1.1745
## log(act_undir_committers + 1) 1.8436    0.5424    1.7504    1.9418
## log(depgraph + 1)       2.1370    0.4679    1.8651    2.4486
## log(depsim + 1)         2.4750    0.4040    1.5801    3.8769
## log(desc + 1)           4.4011    0.2272    3.3340    5.8097
##
## Concordance= 0.834 (se = 0.006 )
## Likelihood ratio test= 2662 on 10 df,  p=<2e-16
## Wald test              = 3326 on 10 df,  p=<2e-16
## Score (logrank) test = 5064 on 10 df,  p=<2e-16
```

Running Survival Models – codeship

```
coxph_codeship <- coxph(Surv(start_interval,
                             end_interval,
                             adopted == 1) ~
  + log(n_commits+1)
  + log(n_watchers+1)
  + (other_badge > 0)
  + log(all_popularity+1)
  + (comp_badge>0)
  + log(act_watchers+1)
  + log(act_undir_committers+1)
  + log(depgraph+1)
  #+ log(depsim+1)
  #+ log(desc+1)
  , data=df_codeship)
summary(coxph_codeship)

## Call:
## coxph(formula = Surv(start_interval, end_interval, adopted ==
##      1) ~ +log(n_commits + 1) + log(n_watchers + 1) + (other_badge >
##      0) + log(all_popularity + 1) + (comp_badge > 0) + log(act_watchers +
##      1) + log(act_undir_committers + 1) + log(depgraph + 1), data = df_codeship)
##
##      n= 19650, number of events= 317
##
##               coef exp(coef) se(coef)      z
## log(n_commits + 1)    0.429385  1.536312  0.046407  9.253
## log(n_watchers + 1) -0.029698  0.970739  0.051077 -0.581
## other_badge > OTRUE    0.543404  1.721859  0.132558  4.099
```

```
## log(all_popularity + 1)      0.040113  1.040929  0.061862  0.648
## comp_badge > OTRUE          -1.916519  0.147118  0.249513 -7.681
## log(act_watchers + 1)       0.003811  1.003819  0.128867  0.030
## log(act_undir_committers + 1) 0.985207  2.678367  0.071487 13.782
## log(depgraph + 1)           1.629057  5.099065  0.375620  4.337
##                               Pr(>|z|)
## log(n_commits + 1)          < 2e-16 ***
## log(n_watchers + 1)         0.561
## other_badge > OTRUE         4.14e-05 ***
## log(all_popularity + 1)     0.517
## comp_badge > OTRUE          1.58e-14 ***
## log(act_watchers + 1)       0.976
## log(act_undir_committers + 1) < 2e-16 ***
## log(depgraph + 1)           1.44e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
##                               exp(coef) exp(-coef) lower .95 upper .95
## log(n_commits + 1)          1.5363      0.6509   1.40274   1.6826
## log(n_watchers + 1)         0.9707      1.0301   0.87826   1.0729
## other_badge > OTRUE         1.7219      0.5808   1.32790   2.2327
## log(all_popularity + 1)     1.0409      0.9607   0.92207   1.1751
## comp_badge > OTRUE          0.1471      6.7973   0.09022   0.2399
## log(act_watchers + 1)       1.0038      0.9962   0.77976   1.2923
## log(act_undir_committers + 1) 2.6784      0.3734   2.32820   3.0812
## log(depgraph + 1)           5.0991      0.1961   2.44209  10.6468
##
## Concordance= 0.849  (se = 0.012 )
## Likelihood ratio test= 586.1  on 8 df,   p=<2e-16
## Wald test              = 795.7  on 8 df,   p=<2e-16
## Score (logrank) test = 1831  on 8 df,   p=<2e-16
```

Running Survival Models – coveralls

```
coxph_coveralls <- coxph(Surv(start_interval,
                              end_interval,
                              adopted == 1) ~
  + log(n_commits+1)
  + log(n_watchers+1)
  + (other_badge > 0)
  + log(all_popularity+1)
  + (comp_badge>0)
  + log(act_watchers+1)
  + log(act_undir_committers+1)
  + log(depgraph+1)
  + log(depsim+1)
  + log(desc+1)
  , data=df_coveralls)
summary(coxph_coveralls)

## Call:
## coxph(formula = Surv(start_interval, end_interval, adopted ==
##      1) ~ +log(n_commits + 1) + log(n_watchers + 1) + (other_badge >
```

```
##      0) + log(all_popularity + 1) + (comp_badge > 0) + log(act_watchers +
##      1) + log(act_undir_committers + 1) + log(depgraph + 1) +
##      log(depsim + 1) + log(desc + 1), data = df_coveralls)
##
##      n= 492812, number of events= 8021
##
##              coef exp(coef) se(coef)      z
## log(n_commits + 1)      0.505328  1.657530  0.010419  48.501
## log(n_watchers + 1)     -0.130159  0.877956  0.009598 -13.560
## other_badge > OTRUE      1.068235  2.910240  0.030261  35.300
## log(all_popularity + 1)  0.127160  1.135599  0.014491   8.775
## comp_badge > OTRUE     -0.094981  0.909390  0.130082  -0.730
## log(act_watchers + 1)   -0.044806  0.956183  0.009251  -4.843
## log(act_undir_committers + 1) 0.518182  1.678973  0.010621  48.791
## log(depgraph + 1)       0.499649  1.648143  0.027533  18.147
## log(depsim + 1)        -0.206924  0.813081  0.066305  -3.121
## log(desc + 1)          0.749356  2.115637  0.052992  14.141
##
##              Pr(>|z|)
## log(n_commits + 1)      < 2e-16 ***
## log(n_watchers + 1)     < 2e-16 ***
## other_badge > OTRUE     < 2e-16 ***
## log(all_popularity + 1) < 2e-16 ***
## comp_badge > OTRUE      0.4653
## log(act_watchers + 1)   1.28e-06 ***
## log(act_undir_committers + 1) < 2e-16 ***
## log(depgraph + 1)      < 2e-16 ***
## log(depsim + 1)        0.0018 **
## log(desc + 1)          < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
##              exp(coef) exp(-coef) lower .95 upper .95
## log(n_commits + 1)      1.6575      0.6033      1.6240      1.6917
## log(n_watchers + 1)     0.8780      1.1390      0.8616      0.8946
## other_badge > OTRUE     2.9102      0.3436      2.7426      3.0881
## log(all_popularity + 1) 1.1356      0.8806      1.1038      1.1683
## comp_badge > OTRUE      0.9094      1.0996      0.7047      1.1735
## log(act_watchers + 1)   0.9562      1.0458      0.9390      0.9737
## log(act_undir_committers + 1) 1.6790      0.5956      1.6444      1.7143
## log(depgraph + 1)       1.6481      0.6067      1.5616      1.7395
## log(depsim + 1)        0.8131      1.2299      0.7140      0.9259
## log(desc + 1)          2.1156      0.4727      1.9069      2.3472
##
## Concordance= 0.8 (se = 0.003 )
## Likelihood ratio test= 9644 on 10 df,  p=<2e-16
## Wald test              = 11208 on 10 df,  p=<2e-16
## Score (logrank) test = 13152 on 10 df,  p=<2e-16
```

Running Survival Models – codeclimate

```
coxph_codeclimate <- coxph(Surv(start_interval,
                                end_interval,
```

```

        adopted == 1) ~
+ log(n_commits+1)
+ log(n_watchers+1)
+ (other_badge > 0)
+ log(all_popularity+1)
+ (comp_badge>0)
+ log(act_watchers+1)
+ log(act_undir_committers+1)
+ log(depgraph+1)
+ log(depsim+1)
  #+ log(desc+1)
  , data=df_codeclimate)
summary(coxph_codeclimate)

## Call:
## coxph(formula = Surv(start_interval, end_interval, adopted ==
##      1) ~ +log(n_commits + 1) + log(n_watchers + 1) + (other_badge >
##      0) + log(all_popularity + 1) + (comp_badge > 0) + log(act_watchers +
##      1) + log(act_undir_committers + 1) + log(depgraph + 1) +
##      log(depsim + 1), data = df_codeclimate)
##
##      n= 210007, number of events= 3306
##
##               coef exp(coef) se(coef)      z Pr(>|z|)
## log(n_commits + 1)      0.53787   1.71235  0.01622 33.167 < 2e-16
## log(n_watchers + 1)     -0.09657   0.90795  0.01478 -6.535 6.36e-11
## other_badge > OTRUE      0.99702   2.71019  0.04755 20.970 < 2e-16
## log(all_popularity + 1)  0.05806   1.05978  0.02103  2.761 0.00575
## comp_badge > OTRUE      0.85648   2.35485  0.07320 11.701 < 2e-16
## log(act_watchers + 1)    -0.16550   0.84747  0.02025 -8.172 3.03e-16
## log(act_undir_committers + 1) 0.61421   1.84820  0.01639 37.482 < 2e-16
## log(depgraph + 1)        0.88879   2.43219  0.05892 15.086 < 2e-16
## log(depsim + 1)         0.14538   1.15648  0.16702  0.870 0.38408
##
## log(n_commits + 1)      ***
## log(n_watchers + 1)      ***
## other_badge > OTRUE      ***
## log(all_popularity + 1)   **
## comp_badge > OTRUE      ***
## log(act_watchers + 1)     ***
## log(act_undir_committers + 1) ***
## log(depgraph + 1)        ***
## log(depsim + 1)
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
##               exp(coef) exp(-coef) lower .95 upper .95
## log(n_commits + 1)      1.7123    0.5840    1.6588    1.7676
## log(n_watchers + 1)      0.9079    1.1014    0.8820    0.9346
## other_badge > OTRUE      2.7102    0.3690    2.4690    2.9749
## log(all_popularity + 1)  1.0598    0.9436    1.0170    1.1044
## comp_badge > OTRUE      2.3549    0.4247    2.0401    2.7181
## log(act_watchers + 1)    0.8475    1.1800    0.8145    0.8818
## log(act_undir_committers + 1) 1.8482    0.5411    1.7898    1.9085

```

```
## log(depgraph + 1)          2.4322      0.4112      2.1669      2.7299
## log(depsim + 1)           1.1565      0.8647      0.8336      1.6044
##
## Concordance= 0.812 (se = 0.005 )
## Likelihood ratio test= 4666 on 9 df, p=<2e-16
## Wald test                = 5994 on 9 df, p=<2e-16
## Score (logrank) test = 8599 on 9 df, p=<2e-16
```

Running Survival Models – codecov

```
coxph_codecov <- coxph(Surv(start_interval,
                             end_interval,
                             adopted == 1) ~
  + log(n_commits+1)
  + log(n_watchers+1)
  + (other_badge > 0)
  + log(all_popularity+1)
  + (comp_badge>0)
  + log(act_watchers+1)
  + log(act_undir_committers+1)
  + log(depgraph+1)
  + log(depsim+1)
  + log(desc+1)
  , data=df_codecov)
summary(coxph_codecov)
```

```
## Call:
## coxph(formula = Surv(start_interval, end_interval, adopted ==
##      1) ~ +log(n_commits + 1) + log(n_watchers + 1) + (other_badge >
##      0) + log(all_popularity + 1) + (comp_badge > 0) + log(act_watchers +
##      1) + log(act_undir_committers + 1) + log(depgraph + 1) +
##      log(depsim + 1) + log(desc + 1), data = df_codecov)
##
##      n= 170932, number of events= 2570
##
##              coef exp(coef) se(coef)      z Pr(>|z|)
## log(n_commits + 1)      0.36135   1.43527  0.01833 19.715 < 2e-16
## log(n_watchers + 1)    -0.05116   0.95012  0.01534 -3.334 0.000855
## other_badge > OTRUE      1.05811   2.88091  0.05751 18.400 < 2e-16
## log(all_popularity + 1)  1.37356   3.94940  0.05781 23.760 < 2e-16
## comp_badge > OTRUE     -0.41080   0.66312  0.17562 -2.339 0.019329
## log(act_watchers + 1)    0.04836   1.04955  0.02162  2.237 0.025282
## log(act_undir_committers + 1) 0.70215   2.01808  0.02250 31.213 < 2e-16
## log(depgraph + 1)       0.51879   1.67999  0.08220  6.311 2.76e-10
## log(depsim + 1)         0.25245   1.28718  0.12496  2.020 0.043363
## log(desc + 1)          0.19536   1.21575  0.04361  4.480 7.46e-06
##
## log(n_commits + 1)      ***
## log(n_watchers + 1)    ***
## other_badge > OTRUE      ***
## log(all_popularity + 1)  ***
## comp_badge > OTRUE       *
## log(act_watchers + 1)    *
```

```
## log(act_undir_committers + 1) ***
## log(depgraph + 1) ***
## log(depsim + 1) *
## log(desc + 1) ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
##
## exp(coef) exp(-coef) lower .95 upper .95
## log(n_commits + 1) 1.4353 0.6967 1.385 1.4878
## log(n_watchers + 1) 0.9501 1.0525 0.922 0.9791
## other_badge > 0TRUE 2.8809 0.3471 2.574 3.2246
## log(all_popularity + 1) 3.9494 0.2532 3.526 4.4232
## comp_badge > 0TRUE 0.6631 1.5080 0.470 0.9356
## log(act_watchers + 1) 1.0496 0.9528 1.006 1.0950
## log(act_undir_committers + 1) 2.0181 0.4955 1.931 2.1091
## log(depgraph + 1) 1.6800 0.5952 1.430 1.9737
## log(depsim + 1) 1.2872 0.7769 1.008 1.6444
## log(desc + 1) 1.2157 0.8225 1.116 1.3242
##
## Concordance= 0.858 (se = 0.004 )
## Likelihood ratio test= 4917 on 10 df, p=<2e-16
## Wald test = 5984 on 10 df, p=<2e-16
## Score (logrank) test = 12793 on 10 df, p=<2e-16
```

Running Survival Models – codacy

```
coxph_codacy <- coxph(Surv(start_interval,
                           end_interval,
                           adopted == 1) ~
  + log(n_commits+1)
  + log(n_watchers+1)
  + (other_badge > 0)
  + log(all_popularity+1)
  #+ (comp_badge>0)
  + log(act_watchers+1)
  + log(act_undir_committers+1)
  + log(depgraph+1)
  + log(depsim+1)
  + log(desc+1)
  , data=df_codacy)
summary(coxph_codacy)

## Call:
## coxph(formula = Surv(start_interval, end_interval, adopted ==
## 1) ~ +log(n_commits + 1) + log(n_watchers + 1) + (other_badge >
## 0) + log(all_popularity + 1) + log(act_watchers + 1) + log(act_undir_committers +
## 1) + log(depgraph + 1) + log(depsim + 1) + log(desc + 1),
## data = df_codacy)
##
## n= 38381, number of events= 596
##
##      coef exp(coef) se(coef)      z
## log(n_commits + 1) 0.56583 1.76091 0.03782 14.962
```

```
## log(n_watchers + 1)          -0.09524  0.90915  0.03205 -2.972
## other_badge > OTRUE          0.90569  2.47364  0.11873  7.628
## log(all_popularity + 1)      0.46434  1.59097  0.07432  6.248
## log(act_watchers + 1)        0.08844  1.09247  0.07672  1.153
## log(act_undir_committers + 1) 1.13418  3.10863  0.06525 17.382
## log(depgraph + 1)           0.75023  2.11749  0.17693  4.240
## log(depsim + 1)             0.41137  1.50889  0.58655  0.701
## log(desc + 1)               5.20430 182.05407  0.41361 12.583
##                               Pr(>|z|)
## log(n_commits + 1)          < 2e-16 ***
## log(n_watchers + 1)         0.00296 **
## other_badge > OTRUE         2.38e-14 ***
## log(all_popularity + 1)     4.15e-10 ***
## log(act_watchers + 1)       0.24901
## log(act_undir_committers + 1) < 2e-16 ***
## log(depgraph + 1)          2.23e-05 ***
## log(depsim + 1)            0.48308
## log(desc + 1)              < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
##                               exp(coef) exp(-coef) lower .95 upper .95
## log(n_commits + 1)          1.7609  0.567888  1.6351  1.8964
## log(n_watchers + 1)         0.9092  1.099925  0.8538  0.9681
## other_badge > OTRUE         2.4736  0.404262  1.9601  3.1218
## log(all_popularity + 1)     1.5910  0.628547  1.3753  1.8404
## log(act_watchers + 1)       1.0925  0.915355  0.9399  1.2697
## log(act_undir_committers + 1) 3.1086  0.321686  2.7354  3.5327
## log(depgraph + 1)          2.1175  0.472258  1.4970  2.9952
## log(depsim + 1)            1.5089  0.662739  0.4780  4.7635
## log(desc + 1)              182.0541  0.005493  80.9348 409.5107
##
## Concordance= 0.852 (se = 0.01 )
## Likelihood ratio test= 1136 on 9 df,  p=<2e-16
## Wald test              = 1464 on 9 df,  p=<2e-16
## Score (logrank) test = 3817 on 9 df,  p=<2e-16
```

```
vif(coxph_codacy)
```

```
## Warning in vif.default(coxph_codacy): No intercept: vifs may not be
## sensible.
```

```
##           log(n_commits + 1)           log(n_watchers + 1)
##           1.254504             1.208324
##           other_badge > 0       log(all_popularity + 1)
##           1.071538             1.058498
##           log(act_watchers + 1) log(act_undir_committers + 1)
##           1.324097             1.564388
##           log(depgraph + 1)       log(depsim + 1)
##           1.418024             1.118240
##           log(desc + 1)
##           1.044503
```

```
anova(coxph_codacy)
```

```
## Analysis of Deviance Table
```



```

## Cox model: response is Surv(start_interval, end_interval, adopted == 1)
## Terms added sequentially (first to last)
##
##               loglik      Chisq Df Pr(>|Chi|)
## NULL                      -4245.2
## log(n_commits + 1)        -4026.1 438.1075 1 < 2.2e-16 ***
## log(n_watchers + 1)       -4020.7  10.8323 1 0.0009975 ***
## other_badge > 0          -3953.5 134.4584 1 < 2.2e-16 ***
## log(all_popularity + 1)   -3911.9  83.0655 1 < 2.2e-16 ***
## log(act_watchers + 1)     -3893.4  37.0126 1 1.174e-09 ***
## log(act_undir_committers + 1) -3726.5 333.9373 1 < 2.2e-16 ***
## log(depgraph + 1)        -3717.8  17.2322 1 3.308e-05 ***
## log(depsim + 1)          -3716.9   1.9475 1 0.1628539
## log(desc + 1)            -3677.0  79.7274 1 < 2.2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```