

Aufgabe4

January 30, 2017

1 Daten Importieren

```
In [2]: import pandas as pd
import numpy as np
```

```
df = pd.read_csv('/media/jascha/530abec2-0ac0-4858-bf4b-242cc3dfc37d/Data/travistorrent_

print "Import ready!"
```

Import ready!

1.0.1 Export der Spalte "gh_project_name" für das github-clone-script

```
In [3]: projects = df['gh_project_name']
```

```
In [12]: unique_projects = projects.unique()
unique_projects
```

```
np.savetxt('unique_projects', unique_projects, fmt='%s',newline='\n')
```

2 Überblick über die Daten verschaffen

```
In [7]: df.shape
```

```
Out[7]: (3702595, 61)
```

```
In [8]: df.head
```

```
Out[8]: <bound method DataFrame.head of
```

		tr_build_id		gh_project_name	gh_is_
0	106060	rails/rails	FALSE	NaN	
1	106060	rails/rails	FALSE	NaN	
2	106060	rails/rails	FALSE	NaN	
3	106060	rails/rails	FALSE	NaN	
4	106060	rails/rails	FALSE	NaN	
5	106060	rails/rails	FALSE	NaN	
6	106060	rails/rails	FALSE	NaN	

7	106060	rails/rails	FALSE	NaN
8	106060	rails/rails	FALSE	NaN
9	106060	rails/rails	FALSE	NaN
10	106060	rails/rails	FALSE	NaN
11	106060	rails/rails	FALSE	NaN
12	106060	rails/rails	FALSE	NaN
13	106060	rails/rails	FALSE	NaN
14	106060	rails/rails	FALSE	NaN
15	106176	rails/rails	FALSE	NaN
16	106176	rails/rails	FALSE	NaN
17	106176	rails/rails	FALSE	NaN
18	106176	rails/rails	FALSE	NaN
19	106176	rails/rails	FALSE	NaN
20	106176	rails/rails	FALSE	NaN
21	106176	rails/rails	FALSE	NaN
22	106176	rails/rails	FALSE	NaN
23	106176	rails/rails	FALSE	NaN
24	106176	rails/rails	FALSE	NaN
25	106176	rails/rails	FALSE	NaN
26	106176	rails/rails	FALSE	NaN
27	106176	rails/rails	FALSE	NaN
28	106176	rails/rails	FALSE	NaN
29	106176	rails/rails	FALSE	NaN
...
3702565	156664398	rails/rails	FALSE	NaN
3702566	156664398	rails/rails	FALSE	NaN
3702567	156664540	apache/storm	FALSE	NaN
3702568	156664540	apache/storm	FALSE	NaN
3702569	156664582	apache/storm	FALSE	NaN
3702570	156664582	apache/storm	FALSE	NaN
3702571	156664582	apache/storm	FALSE	NaN
3702572	156664582	apache/storm	FALSE	NaN
3702573	156664590	apache/storm	FALSE	NaN
3702574	156664590	apache/storm	FALSE	NaN
3702575	156664590	apache/storm	FALSE	NaN
3702576	156664590	apache/storm	FALSE	NaN
3702577	156667224	projectblacklight/blacklight	FALSE	NaN
3702578	156667224	projectblacklight/blacklight	FALSE	NaN
3702579	156667224	projectblacklight/blacklight	FALSE	NaN
3702580	156667224	projectblacklight/blacklight	FALSE	NaN
3702581	156667627	projecthydra/sufia	FALSE	NaN
3702582	156667627	projecthydra/sufia	FALSE	NaN
3702583	156668690	yegor256/rultor	FALSE	NaN
3702584	156669587	owncloud/android	FALSE	NaN
3702585	156670030	projectblacklight/blacklight	FALSE	NaN
3702586	156670030	projectblacklight/blacklight	FALSE	NaN
3702587	156670030	projectblacklight/blacklight	FALSE	NaN
3702588	156670030	projectblacklight/blacklight	FALSE	NaN

3702589	156671240	projecthydra/sufia	FALSE	NaN
3702590	156671240	projecthydra/sufia	FALSE	NaN
3702591	156675548	svenfuchs/rails-i18n	FALSE	NaN
3702592	156675548	svenfuchs/rails-i18n	FALSE	NaN
3702593	156676294	google/closure-compiler	FALSE	NaN
3702594	156676294	google/closure-compiler	FALSE	NaN

	gh_pull_req_num	gh_lang	git_merged_with	\
0	NaN	ruby	NaN	
1	NaN	ruby	NaN	
2	NaN	ruby	NaN	
3	NaN	ruby	NaN	
4	NaN	ruby	NaN	
5	NaN	ruby	NaN	
6	NaN	ruby	NaN	
7	NaN	ruby	NaN	
8	NaN	ruby	NaN	
9	NaN	ruby	NaN	
10	NaN	ruby	NaN	
11	NaN	ruby	NaN	
12	NaN	ruby	NaN	
13	NaN	ruby	NaN	
14	NaN	ruby	NaN	
15	NaN	ruby	NaN	
16	NaN	ruby	NaN	
17	NaN	ruby	NaN	
18	NaN	ruby	NaN	
19	NaN	ruby	NaN	
20	NaN	ruby	NaN	
21	NaN	ruby	NaN	
22	NaN	ruby	NaN	
23	NaN	ruby	NaN	
24	NaN	ruby	NaN	
25	NaN	ruby	NaN	
26	NaN	ruby	NaN	
27	NaN	ruby	NaN	
28	NaN	ruby	NaN	
29	NaN	ruby	NaN	
...	
3702565	NaN	ruby	NaN	
3702566	NaN	ruby	NaN	
3702567	NaN	java	NaN	
3702568	NaN	java	NaN	
3702569	NaN	java	NaN	
3702570	NaN	java	NaN	
3702571	NaN	java	NaN	
3702572	NaN	java	NaN	
3702573	NaN	java	NaN	

3702574	NaN	java	NaN
3702575	NaN	java	NaN
3702576	NaN	java	NaN
3702577	NaN	ruby	NaN
3702578	NaN	ruby	NaN
3702579	NaN	ruby	NaN
3702580	NaN	ruby	NaN
3702581	NaN	ruby	NaN
3702582	NaN	ruby	NaN
3702583	NaN	java	NaN
3702584	NaN	java	NaN
3702585	NaN	ruby	NaN
3702586	NaN	ruby	NaN
3702587	NaN	ruby	NaN
3702588	NaN	ruby	NaN
3702589	NaN	ruby	NaN
3702590	NaN	ruby	NaN
3702591	NaN	ruby	NaN
3702592	NaN	ruby	NaN
3702593	NaN	java	NaN
3702594	NaN	java	NaN

	git_branch	gh_num_commits_in_push	\
0	3-1-stable		NaN
1	3-1-stable		NaN
2	3-1-stable		NaN
3	3-1-stable		NaN
4	3-1-stable		NaN
5	3-1-stable		NaN
6	3-1-stable		NaN
7	3-1-stable		NaN
8	3-1-stable		NaN
9	3-1-stable		NaN
10	3-1-stable		NaN
11	3-1-stable		NaN
12	3-1-stable		NaN
13	3-1-stable		NaN
14	3-1-stable		NaN
15	master		NaN
16	master		NaN
17	master		NaN
18	master		NaN
19	master		NaN
20	master		NaN
21	master		NaN
22	master		NaN
23	master		NaN
24	master		NaN

25	master	NaN
26	master	NaN
27	master	NaN
28	master	NaN
29	master	NaN
...
3702565	master	NaN
3702566	master	NaN
3702567	master	NaN
3702568	master	NaN
3702569	1.0.x-branch	NaN
3702570	1.0.x-branch	NaN
3702571	1.0.x-branch	NaN
3702572	1.0.x-branch	NaN
3702573	1.x-branch	NaN
3702574	1.x-branch	NaN
3702575	1.x-branch	NaN
3702576	1.x-branch	NaN
3702577	render_filter_element_strings	NaN
3702578	render_filter_element_strings	NaN
3702579	render_filter_element_strings	NaN
3702580	render_filter_element_strings	NaN
3702581	flipflop	NaN
3702582	flipflop	NaN
3702583	master	NaN
3702584	master	NaN
3702585	master	NaN
3702586	master	NaN
3702587	master	NaN
3702588	master	NaN
3702589	master	NaN
3702590	master	NaN
3702591	master	NaN
3702592	master	NaN
3702593	master	NaN
3702594	master	NaN

	gh_commits_in_push	...	tr_log_frameworks \
0	NaN	...	NaN
1	NaN	...	NaN
2	NaN	...	NaN
3	NaN	...	NaN
4	NaN	...	NaN
5	NaN	...	NaN
6	NaN	...	NaN
7	NaN	...	NaN
8	NaN	...	NaN
9	NaN	...	NaN

10	NaN	...	NaN
11	NaN	...	NaN
12	NaN	...	NaN
13	NaN	...	NaN
14	NaN	...	NaN
15	NaN	...	NaN
16	NaN	...	NaN
17	NaN	...	NaN
18	NaN	...	NaN
19	NaN	...	NaN
20	NaN	...	NaN
21	NaN	...	NaN
22	NaN	...	NaN
23	NaN	...	NaN
24	NaN	...	NaN
25	NaN	...	NaN
26	NaN	...	NaN
27	NaN	...	NaN
28	NaN	...	NaN
29	NaN	...	NaN
...
3702565	NaN	...	NaN
3702566	NaN	...	testunit
3702567	NaN	...	NaN
3702568	NaN	...	NaN
3702569	NaN	...	NaN
3702570	NaN	...	NaN
3702571	NaN	...	NaN
3702572	NaN	...	NaN
3702573	NaN	...	NaN
3702574	NaN	...	NaN
3702575	NaN	...	NaN
3702576	NaN	...	NaN
3702577	NaN	...	rspec
3702578	NaN	...	rspec
3702579	NaN	...	rspec
3702580	NaN	...	NaN
3702581	NaN	...	rspec
3702582	NaN	...	rspec
3702583	NaN	...	NaN
3702584	NaN	...	NaN
3702585	NaN	...	NaN
3702586	NaN	...	rspec
3702587	NaN	...	rspec
3702588	NaN	...	NaN
3702589	NaN	...	rspec
3702590	NaN	...	rspec
3702591	NaN	...	rspec

3702592	NaN	...	rspec
3702593	NaN	...	junit
3702594	NaN	...	junit

	tr_log_bool_tests_ran	tr_log_bool_tests_failed	tr_log_num_tests_ok	\
0	FALSE	NaN	NaN	
1	FALSE	NaN	NaN	
2	FALSE	NaN	NaN	
3	FALSE	NaN	NaN	
4	FALSE	NaN	NaN	
5	FALSE	NaN	NaN	
6	FALSE	NaN	NaN	
7	FALSE	NaN	NaN	
8	FALSE	NaN	NaN	
9	FALSE	NaN	NaN	
10	FALSE	NaN	NaN	
11	FALSE	NaN	NaN	
12	FALSE	NaN	NaN	
13	FALSE	NaN	NaN	
14	FALSE	NaN	NaN	
15	FALSE	NaN	NaN	
16	FALSE	NaN	NaN	
17	FALSE	NaN	NaN	
18	FALSE	NaN	NaN	
19	FALSE	NaN	NaN	
20	FALSE	NaN	NaN	
21	FALSE	NaN	NaN	
22	FALSE	NaN	NaN	
23	FALSE	NaN	NaN	
24	FALSE	NaN	NaN	
25	FALSE	NaN	NaN	
26	FALSE	NaN	NaN	
27	FALSE	NaN	NaN	
28	FALSE	NaN	NaN	
29	FALSE	NaN	NaN	
...	
3702565	FALSE	NaN	NaN	
3702566	TRUE	FALSE	6076	
3702567	FALSE	NaN	NaN	
3702568	FALSE	NaN	NaN	
3702569	FALSE	NaN	NaN	
3702570	FALSE	NaN	NaN	
3702571	FALSE	NaN	NaN	
3702572	FALSE	NaN	NaN	
3702573	FALSE	NaN	NaN	
3702574	FALSE	NaN	NaN	
3702575	FALSE	NaN	NaN	
3702576	FALSE	NaN	NaN	

3702577	TRUE	FALSE	1111
3702578	TRUE	FALSE	1111
3702579	TRUE	FALSE	1111
3702580	FALSE	NaN	NaN
3702581	TRUE	FALSE	1086
3702582	TRUE	FALSE	1086
3702583	FALSE	NaN	NaN
3702584	FALSE	NaN	NaN
3702585	FALSE	NaN	NaN
3702586	TRUE	FALSE	1111
3702587	TRUE	FALSE	1111
3702588	FALSE	NaN	NaN
3702589	TRUE	FALSE	1086
3702590	TRUE	FALSE	1086
3702591	TRUE	FALSE	2734
3702592	TRUE	FALSE	2734
3702593	TRUE	FALSE	10975
3702594	TRUE	FALSE	10975

	tr_log_num_tests_failed	tr_log_num_tests_run	tr_log_num_tests_skipped	\
0	NaN	NaN	NaN	
1	NaN	NaN	NaN	
2	NaN	NaN	NaN	
3	NaN	NaN	NaN	
4	NaN	NaN	NaN	
5	NaN	NaN	NaN	
6	NaN	NaN	NaN	
7	NaN	NaN	NaN	
8	NaN	NaN	NaN	
9	NaN	NaN	NaN	
10	NaN	NaN	NaN	
11	NaN	NaN	NaN	
12	NaN	NaN	NaN	
13	NaN	NaN	NaN	
14	NaN	NaN	NaN	
15	NaN	NaN	NaN	
16	NaN	NaN	NaN	
17	NaN	NaN	NaN	
18	NaN	NaN	NaN	
19	NaN	NaN	NaN	
20	NaN	NaN	NaN	
21	NaN	NaN	NaN	
22	NaN	NaN	NaN	
23	NaN	NaN	NaN	
24	NaN	NaN	NaN	
25	NaN	NaN	NaN	
26	NaN	NaN	NaN	
27	NaN	NaN	NaN	

28	NaN	NaN	NaN
29	NaN	NaN	NaN
...
3702565	NaN	NaN	NaN
3702566	0	6076	2
3702567	NaN	NaN	NaN
3702568	NaN	NaN	NaN
3702569	NaN	NaN	NaN
3702570	NaN	NaN	NaN
3702571	NaN	NaN	NaN
3702572	NaN	NaN	NaN
3702573	NaN	NaN	NaN
3702574	NaN	NaN	NaN
3702575	NaN	NaN	NaN
3702576	NaN	NaN	NaN
3702577	0	1111	10
3702578	0	1111	10
3702579	0	1111	10
3702580	NaN	NaN	NaN
3702581	0	1086	8
3702582	0	1086	8
3702583	NaN	NaN	NaN
3702584	NaN	NaN	NaN
3702585	NaN	NaN	NaN
3702586	0	1111	10
3702587	0	1111	10
3702588	NaN	NaN	NaN
3702589	0	1086	8
3702590	0	1086	8
3702591	0	2734	0
3702592	0	2734	0
3702593	0	10975	0
3702594	0	10975	0

	tr_log_tests_failed	tr_log_testduration	tr_log_buildduration
0	NaN	NaN	NaN
1	NaN	NaN	NaN
2	NaN	NaN	NaN
3	NaN	NaN	NaN
4	NaN	NaN	NaN
5	NaN	NaN	NaN
6	NaN	NaN	NaN
7	NaN	NaN	NaN
8	NaN	NaN	NaN
9	NaN	NaN	NaN
10	NaN	NaN	NaN
11	NaN	NaN	NaN
12	NaN	NaN	NaN

13	NaN	NaN	NaN
14	NaN	NaN	NaN
15	NaN	NaN	NaN
16	NaN	NaN	NaN
17	NaN	NaN	NaN
18	NaN	NaN	NaN
19	NaN	NaN	NaN
20	NaN	NaN	NaN
21	NaN	NaN	NaN
22	NaN	NaN	NaN
23	NaN	NaN	NaN
24	NaN	NaN	NaN
25	NaN	NaN	NaN
26	NaN	NaN	NaN
27	NaN	NaN	NaN
28	NaN	NaN	NaN
29	NaN	NaN	NaN
...
3702565	NaN	NaN	NaN
3702566	NaN	19.01	NaN
3702567	NaN	NaN	NaN
3702568	NaN	NaN	NaN
3702569	NaN	NaN	248
3702570	NaN	NaN	NaN
3702571	NaN	NaN	NaN
3702572	NaN	NaN	NaN
3702573	NaN	NaN	NaN
3702574	NaN	NaN	NaN
3702575	NaN	NaN	NaN
3702576	NaN	NaN	NaN
3702577	NaN	6.26	NaN
3702578	NaN	2.49	NaN
3702579	NaN	41.61	NaN
3702580	NaN	NaN	NaN
3702581	NaN	6.67	NaN
3702582	NaN	6.77	NaN
3702583	NaN	NaN	111
3702584	NaN	NaN	NaN
3702585	NaN	NaN	NaN
3702586	NaN	51.63	NaN
3702587	NaN	31.68	NaN
3702588	NaN	NaN	NaN
3702589	NaN	9.61	NaN
3702590	NaN	6.96	NaN
3702591	NaN	8.5	NaN
3702592	NaN	10.18	NaN
3702593	NaN	89.5300000000001	183
3702594	NaN	157.7	313

[3702595 rows x 61 columns]>

In [9]: `list(df)`

Out[9]: ['tr_build_id',
'gh_project_name',
'gh_is_pr',
'gh_pr_created_at',
'gh_pull_req_num',
'gh_lang',
'git_merged_with',
'git_branch',
'gh_num_commits_in_push',
'gh_commits_in_push',
'git_prev_commit_resolution_status',
'git_prev_built_commit',
'tr_prev_build',
'gh_first_commit_created_at',
'gh_team_size',
'git_all_built_commits',
'git_num_all_built_commits',
'git_trigger_commit',
'tr_virtual_merged_into',
'tr_original_commit',
'gh_num_issue_comments',
'gh_num_commit_comments',
'gh_num_pr_comments',
'git_diff_src_churn',
'git_diff_test_churn',
'gh_diff_files_added',
'gh_diff_files_deleted',
'gh_diff_files_modified',
'gh_diff_tests_added',
'gh_diff_tests_deleted',
'gh_diff_src_files',
'gh_diff_doc_files',
'gh_diff_other_files',
'gh_num_commits_on_files_touched',
'gh_sloc',
'gh_test_lines_per_kloc',
'gh_test_cases_per_kloc',
'gh_asserts_cases_per_kloc',
'gh_by_core_team_member',
'gh_description_complexity',
'gh_pushed_at',
'gh_build_started_at',
'tr_status',

```

'tr_duration',
'tr_jobs',
'tr_build_number',
'tr_job_id',
'tr_log_lan',
'tr_log_status',
'tr_log_setup_time',
'tr_log_analyzer',
'tr_log_frameworks',
'tr_log_bool_tests_ran',
'tr_log_bool_tests_failed',
'tr_log_num_tests_ok',
'tr_log_num_tests_failed',
'tr_log_num_tests_run',
'tr_log_num_tests_skipped',
'tr_log_tests_failed',
'tr_log_testduration',
'tr_log_buildduration']

```

3 Kleinste und grösSte Projekte herausfinden (nach # Builds)

```

In [19]: agg = df.groupby('gh_project_name').size()
         agg.sort_values()

```

```

Out[19]: gh_project_name
boxen/puppet-osx                2
calabash/calabash-ios           4
carymrobbins/intellij-haskforce 11
code4craft/webmagic            20
emilsjolander/sprinkles         20
bitsofproof/supernode          32
cantino/reckon                 32
goodow/realtime-store          34
boxen/boxen                    39
bugsnag/bugsnag-ruby           48
castorflex/SmoothProgressBar   53
caelum/tubaina                 56
clayallsopp/formotion          63
pedrovg/DraggablePanel         67
twitter/cloudhopper-smpp       69
boxen/puppet-boxen            70
tjerkw/Android-SlideExpandableListView 71
infOrmer/blanket               72
mdxp/nodejs-cookbook           75
codecation/trailmix            76
daimajia/AndroidImageSlider    78
JakeWharton/DiskLruCache       79

```

pbyrne/cron-checker	80
cyu/rack-cors	82
lookfirst/sardine	86
mpatric/mp3agic	87
rockerhieu/emojicon	88
paulasmuth/recommendify	88
ralfstx/minimal-json	94
kevinsawicki/http-request	94
...	
openSUSE/open-build-service	15161
honeybadger-io/honeybadger-ruby	16060
celluloid/celluloid	16256
killbill/killbill	18120
opal/opal	18523
rspec/rspec-mocks	18542
sferik/rails_admin	18733
rubygems/rubygems	18786
spree/spree	19421
neo4jrb/neo4j	19525
ruboto/ruboto	20240
plataformatec/devise	21194
toptal/chewy	21263
fog/fog	21652
rapid7/metasploit-framework	23743
amatsuda/kaminari	24222
rails-api/active_model_serializers	25057
refinery/refinerycms	26146
activerecord-hackery/ransack	28670
rspec/rspec-core	29722
SonarSource/sonarqube	30142
diaspora/diaspora	32393
ruby/ruby	34409
slim-template/slim	43819
rspec/rspec-rails	44397
rollbar/rollbar-gem	49968
jruby/activerecord-jdbc-adapter	51585
opf/openproject	54992
jruby/jruby	271748
rails/rails	471417
dtype: int64	

```
In [4]: df['gh_build_started_at'].head
```

```
Out[4]: <bound method Series.head of 0      2011-08-29 00:25:49
1      2011-08-29 00:25:49
2      2011-08-29 00:25:49
3      2011-08-29 00:25:49
4      2011-08-29 00:25:49
```

5	2011-08-29 00:25:49
6	2011-08-29 00:25:49
7	2011-08-29 00:25:49
8	2011-08-29 00:25:49
9	2011-08-29 00:25:49
10	2011-08-29 00:25:49
11	2011-08-29 00:25:49
12	2011-08-29 00:25:49
13	2011-08-29 00:25:49
14	2011-08-29 00:25:49
15	2011-08-29 00:13:21
16	2011-08-29 00:13:21
17	2011-08-29 00:13:21
18	2011-08-29 00:13:21
19	2011-08-29 00:13:21
20	2011-08-29 00:13:21
21	2011-08-29 00:13:21
22	2011-08-29 00:13:21
23	2011-08-29 00:13:21
24	2011-08-29 00:13:21
25	2011-08-29 00:13:21
26	2011-08-29 00:13:21
27	2011-08-29 00:13:21
28	2011-08-29 00:13:21
29	2011-08-29 00:13:21
	...
3702565	2016-08-31 22:42:07
3702566	2016-08-31 22:42:07
3702567	2016-08-31 23:57:04
3702568	2016-08-31 23:57:04
3702569	2016-08-31 23:58:14
3702570	2016-08-31 23:58:14
3702571	2016-08-31 23:58:14
3702572	2016-08-31 23:58:14
3702573	2016-08-31 23:59:01
3702574	2016-08-31 23:59:01
3702575	2016-08-31 23:59:01
3702576	2016-08-31 23:59:01
3702577	2016-08-31 22:59:46
3702578	2016-08-31 22:59:46
3702579	2016-08-31 22:59:46
3702580	2016-08-31 22:59:46
3702581	2016-08-31 23:01:31
3702582	2016-08-31 23:01:31
3702583	2016-08-31 23:17:40
3702584	2016-08-31 23:14:07
3702585	2016-08-31 23:17:25
3702586	2016-08-31 23:17:25

```

3702587    2016-08-31 23:17:25
3702588    2016-08-31 23:17:25
3702589    2016-08-31 23:25:11
3702590    2016-08-31 23:25:11
3702591    2016-08-31 23:55:57
3702592    2016-08-31 23:55:57
3702593    2016-08-31 23:59:27
3702594    2016-08-31 23:59:27
Name: gh_build_started_at, dtype: object>

```

4 Erster TravisCI Build 08/2011, letzter (in der CSV Datei) 08/2016

5 Nun die Buildzeiten von TravisCI der drei grössten Projekte ausgeben

```

In [30]: hd = df[df['gh_project_name'] == 'rails/rails']
          hd_sort = hd.sort_values(['gh_build_started_at'])
          hd_sort[['gh_project_name', 'gh_build_started_at']]

```

```

Out[30]:
          gh_project_name  gh_build_started_at
25          rails/rails  2011-08-29 00:13:21
29          rails/rails  2011-08-29 00:13:21
28          rails/rails  2011-08-29 00:13:21
27          rails/rails  2011-08-29 00:13:21
26          rails/rails  2011-08-29 00:13:21
24          rails/rails  2011-08-29 00:13:21
23          rails/rails  2011-08-29 00:13:21
22          rails/rails  2011-08-29 00:13:21
21          rails/rails  2011-08-29 00:13:21
20          rails/rails  2011-08-29 00:13:21
19          rails/rails  2011-08-29 00:13:21
18          rails/rails  2011-08-29 00:13:21
17          rails/rails  2011-08-29 00:13:21
16          rails/rails  2011-08-29 00:13:21
15          rails/rails  2011-08-29 00:13:21
0          rails/rails  2011-08-29 00:25:49
13          rails/rails  2011-08-29 00:25:49
14          rails/rails  2011-08-29 00:25:49
1          rails/rails  2011-08-29 00:25:49
2          rails/rails  2011-08-29 00:25:49
3          rails/rails  2011-08-29 00:25:49
5          rails/rails  2011-08-29 00:25:49
6          rails/rails  2011-08-29 00:25:49
4          rails/rails  2011-08-29 00:25:49
8          rails/rails  2011-08-29 00:25:49
9          rails/rails  2011-08-29 00:25:49
10         rails/rails  2011-08-29 00:25:49

```

11	rails/rails	2011-08-29 00:25:49
12	rails/rails	2011-08-29 00:25:49
7	rails/rails	2011-08-29 00:25:49
...
3702559	rails/rails	2016-08-31 22:42:07
3702560	rails/rails	2016-08-31 22:42:07
3702561	rails/rails	2016-08-31 22:42:07
3702562	rails/rails	2016-08-31 22:42:07
3702563	rails/rails	2016-08-31 22:42:07
3702547	rails/rails	2016-08-31 22:42:07
3702556	rails/rails	2016-08-31 22:42:07
3702546	rails/rails	2016-08-31 22:42:07
3702536	rails/rails	2016-08-31 22:42:07
3702544	rails/rails	2016-08-31 22:42:07
3702565	rails/rails	2016-08-31 22:42:07
3702526	rails/rails	2016-08-31 22:42:07
3702527	rails/rails	2016-08-31 22:42:07
3702528	rails/rails	2016-08-31 22:42:07
3702529	rails/rails	2016-08-31 22:42:07
3702530	rails/rails	2016-08-31 22:42:07
3702531	rails/rails	2016-08-31 22:42:07
3702532	rails/rails	2016-08-31 22:42:07
3702545	rails/rails	2016-08-31 22:42:07
3702533	rails/rails	2016-08-31 22:42:07
3702535	rails/rails	2016-08-31 22:42:07
3702537	rails/rails	2016-08-31 22:42:07
3702538	rails/rails	2016-08-31 22:42:07
3702539	rails/rails	2016-08-31 22:42:07
3702540	rails/rails	2016-08-31 22:42:07
3702541	rails/rails	2016-08-31 22:42:07
3702542	rails/rails	2016-08-31 22:42:07
3702543	rails/rails	2016-08-31 22:42:07
3702534	rails/rails	2016-08-31 22:42:07
3702566	rails/rails	2016-08-31 22:42:07

[471417 rows x 2 columns]

```
In [5]: hd2 = df[df['gh_project_name'] == 'opf/openproject']
hd2_sort = hd2.sort_values(['gh_build_started_at'])
hd2_sort[['gh_project_name', 'gh_build_started_at']]
```

```
Out[5]:
```

	gh_project_name	gh_build_started_at
573868	opf/openproject	2013-05-07 22:59:16
573869	opf/openproject	2013-05-07 22:59:16
574553	opf/openproject	2013-05-08 09:32:34
574554	opf/openproject	2013-05-08 09:32:34
574626	opf/openproject	2013-05-08 11:11:17
574627	opf/openproject	2013-05-08 11:11:17

574630	opf/openproject	2013-05-08	11:11:55
574631	opf/openproject	2013-05-08	11:11:55
575142	opf/openproject	2013-05-08	15:17:46
575143	opf/openproject	2013-05-08	15:17:46
575204	opf/openproject	2013-05-08	15:31:32
575205	opf/openproject	2013-05-08	15:31:32
575245	opf/openproject	2013-05-08	16:19:59
575246	opf/openproject	2013-05-08	16:19:59
577960	opf/openproject	2013-05-10	09:47:42
577959	opf/openproject	2013-05-10	09:47:42
577986	opf/openproject	2013-05-10	10:07:02
578007	opf/openproject	2013-05-10	10:55:14
578008	opf/openproject	2013-05-10	10:55:14
578218	opf/openproject	2013-05-10	13:58:13
578219	opf/openproject	2013-05-10	13:58:13
582909	opf/openproject	2013-05-13	15:17:55
582928	opf/openproject	2013-05-13	15:27:27
582929	opf/openproject	2013-05-13	15:27:27
584644	opf/openproject	2013-05-14	12:29:15
584645	opf/openproject	2013-05-14	12:29:15
584717	opf/openproject	2013-05-14	13:34:06
584718	opf/openproject	2013-05-14	13:34:06
584753	opf/openproject	2013-05-14	13:51:28
584752	opf/openproject	2013-05-14	13:51:28
...
2740100	opf/openproject	2015-08-07	15:30:09
2740101	opf/openproject	2015-08-07	15:30:09
2740102	opf/openproject	2015-08-07	15:30:09
2740103	opf/openproject	2015-08-07	15:30:09
2740104	opf/openproject	2015-08-07	15:30:09
2740097	opf/openproject	2015-08-07	15:30:09
2739716	opf/openproject	2015-08-07	15:59:23
2739715	opf/openproject	2015-08-07	15:59:23
2739709	opf/openproject	2015-08-07	15:59:23
2739710	opf/openproject	2015-08-07	15:59:23
2739711	opf/openproject	2015-08-07	15:59:23
2739712	opf/openproject	2015-08-07	15:59:23
2739713	opf/openproject	2015-08-07	15:59:23
2739714	opf/openproject	2015-08-07	15:59:23
2740296	opf/openproject	2015-08-07	16:34:32
2740295	opf/openproject	2015-08-07	16:34:32
2740294	opf/openproject	2015-08-07	16:34:32
2740293	opf/openproject	2015-08-07	16:34:32
2740291	opf/openproject	2015-08-07	16:34:32
2740290	opf/openproject	2015-08-07	16:34:32
2740289	opf/openproject	2015-08-07	16:34:32
2740292	opf/openproject	2015-08-07	16:34:32
2742812	opf/openproject	2015-08-08	18:55:19

```

2742813  opf/openproject  2015-08-08 18:55:19
2742811  opf/openproject  2015-08-08 18:55:19
2742814  opf/openproject  2015-08-08 18:55:19
2742809  opf/openproject  2015-08-08 18:55:19
2742808  opf/openproject  2015-08-08 18:55:19
2742810  opf/openproject  2015-08-08 18:55:19
2742815  opf/openproject  2015-08-08 18:55:19

```

```
[54992 rows x 2 columns]
```

```

In [6]: hd3 = df[df['gh_project_name'] == 'jruby/jruby']
hd3_sort = hd3.sort_values(['gh_build_started_at'])
hd3_sort[['gh_project_name', 'gh_build_started_at']]

```

```

Out [6]:
      gh_project_name  gh_build_started_at
126611      jruby/jruby  2012-03-15 02:07:19
126612      jruby/jruby  2012-03-15 02:11:35
126613      jruby/jruby  2012-03-15 02:17:06
126614      jruby/jruby  2012-03-15 02:19:01
126642      jruby/jruby  2012-03-15 02:39:30
126719      jruby/jruby  2012-03-15 03:21:01
126746      jruby/jruby  2012-03-15 03:50:30
126748      jruby/jruby  2012-03-15 04:32:15
126756      jruby/jruby  2012-03-15 04:44:32
126758      jruby/jruby  2012-03-15 04:59:10
126787      jruby/jruby  2012-03-15 05:19:35
126788      jruby/jruby  2012-03-15 05:20:25
126811      jruby/jruby  2012-03-15 06:19:18
127003      jruby/jruby  2012-03-15 14:07:23
127041      jruby/jruby  2012-03-15 14:27:14
127047      jruby/jruby  2012-03-15 14:56:24
127143      jruby/jruby  2012-03-15 17:29:46
127335      jruby/jruby  2012-03-15 21:06:02
127342      jruby/jruby  2012-03-15 21:12:49
127365      jruby/jruby  2012-03-15 21:57:42
127428      jruby/jruby  2012-03-15 22:22:40
127700      jruby/jruby  2012-03-16 08:45:40
128378      jruby/jruby  2012-03-16 19:54:39
128409      jruby/jruby  2012-03-16 20:15:05
128535      jruby/jruby  2012-03-17 00:29:50
129866      jruby/jruby  2012-03-19 04:49:50
129867      jruby/jruby  2012-03-19 04:53:21
129868      jruby/jruby  2012-03-19 04:55:17
130517      jruby/jruby  2012-03-19 21:30:26
130518      jruby/jruby  2012-03-19 21:32:33
...
3702370      jruby/jruby  2016-08-31 21:33:21
3702439      jruby/jruby  2016-08-31 23:12:26

```

3702436	jruby/jruby	2016-08-31	23:12:26
3702437	jruby/jruby	2016-08-31	23:12:26
3702438	jruby/jruby	2016-08-31	23:12:26
3702440	jruby/jruby	2016-08-31	23:12:26
3702446	jruby/jruby	2016-08-31	23:12:26
3702442	jruby/jruby	2016-08-31	23:12:26
3702443	jruby/jruby	2016-08-31	23:12:26
3702444	jruby/jruby	2016-08-31	23:12:26
3702445	jruby/jruby	2016-08-31	23:12:26
3702435	jruby/jruby	2016-08-31	23:12:26
3702441	jruby/jruby	2016-08-31	23:12:26
3702434	jruby/jruby	2016-08-31	23:12:26
3702427	jruby/jruby	2016-08-31	23:12:26
3702432	jruby/jruby	2016-08-31	23:12:26
3702431	jruby/jruby	2016-08-31	23:12:26
3702430	jruby/jruby	2016-08-31	23:12:26
3702429	jruby/jruby	2016-08-31	23:12:26
3702428	jruby/jruby	2016-08-31	23:12:26
3702426	jruby/jruby	2016-08-31	23:12:26
3702425	jruby/jruby	2016-08-31	23:12:26
3702424	jruby/jruby	2016-08-31	23:12:26
3702423	jruby/jruby	2016-08-31	23:12:26
3702422	jruby/jruby	2016-08-31	23:12:26
3702421	jruby/jruby	2016-08-31	23:12:26
3702420	jruby/jruby	2016-08-31	23:12:26
3702447	jruby/jruby	2016-08-31	23:12:26
3702433	jruby/jruby	2016-08-31	23:12:26
3702448	jruby/jruby	2016-08-31	23:12:26

[271748 rows x 2 columns]

Rails war das erste Projekt, welches TravisCI genutzt hat. Die Datensätze sind von 08/2011 bis 08/2011 vorhanden. jruby hat erst 03/2012 angefangen, auch bis heute. OpenProject startete 05/2013 und beendete die Nutzung von TravisCI 08/2015.