task_management

October 13, 2019

1 Task analysis

Goals:

- Measurement of all tasks created
- Differentiate between unique and double send tasks
- (How many tasks are allready created by workerbase System)
- (How long are workers in general busy out of the 480m)

Current active apps: * - Quality measurements * Machine interruptions * (Trolley change)

1.0.1 Example datapoint:

title

1.0.2 Flattened:

title

1.0.3 Other File

title

1.0.4 Cleaning - Machine Interruptions

First read the data and bring it in the correct schema:

[181]:
```
import pandas as pd
import numpy as np
import cufflinks as cf
import plotly.offline as pyo
import seaborn as sns
import matplotlib.pyplot as plt
```

```
pyo.init_notebook_mode()
cf.go_offline()
cf.set_config_file(offline=False, world_readable=True)

data_mi_history = pd.read_json('5d44361a5a5cbc00067a5f41.taskhistories.json')
data_mi = pd.read_json('5d44361a5a5cbc00067a5f41.tasks.json')

data_mi
```

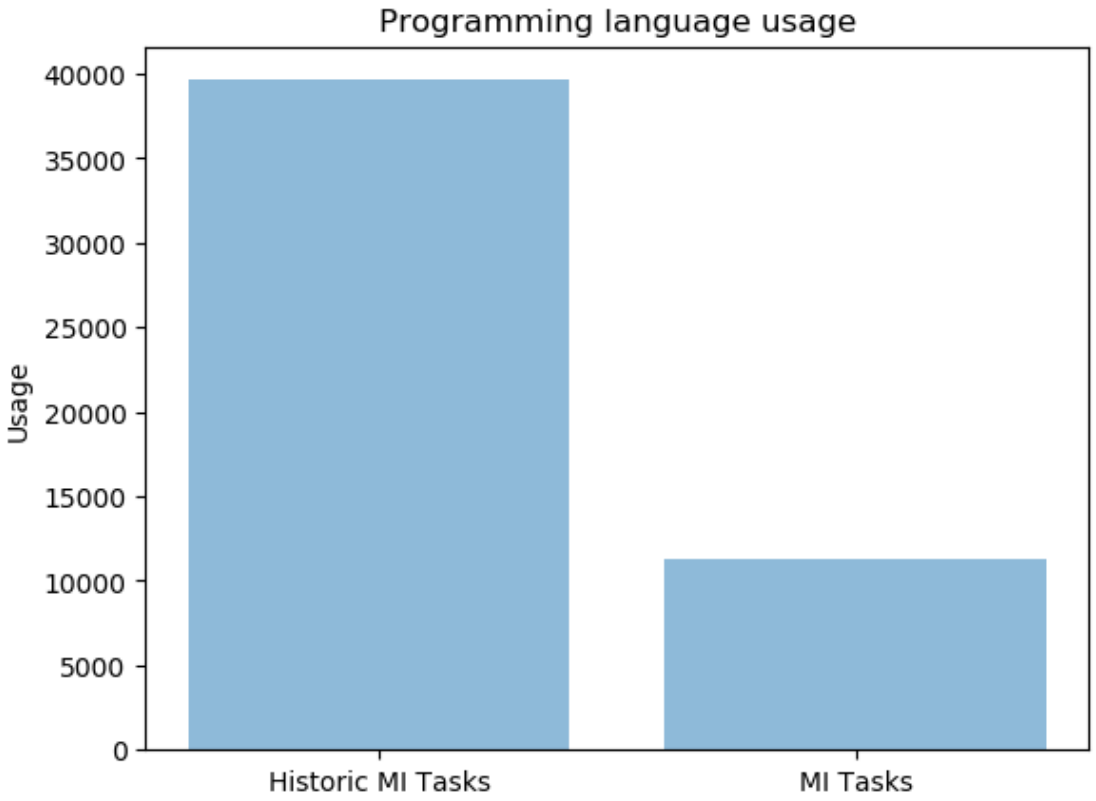
[181]:

```
                                        _id  isApp source status  archived  \
0      {'$oid': '5d4d80813a007a0006292b38'}  False   rule   done      True
1      {'$oid': '5d4d80813a007a0006292b3a'}  False   rule   done      True
2      {'$oid': '5d4d80813a007a0006292b3d'}  False   rule   done      True
3      {'$oid': '5d4d80813a007a0006292b40'}  False   rule   done      True
4      {'$oid': '5d4d80813a007a0006292b44'}  False   rule   done      True
...                                     ...    ...    ...    ...       ...
11257  {'$oid': '5d9d92b4e59c620007b76737'}  False   rule   done     False
11258  {'$oid': '5d9d92d2e59c620007b7673e'}  False   rule   done     False
11259  {'$oid': '5d9d92fbe59c620007b76749'}  False   rule   done     False
11260  {'$oid': '5d9d9318e59c620007b7674e'}  False   rule   done     False
11261  {'$oid': '5d9d9680e59c620007b768dd'}  False   rule   done     False

       deleted                                              payload  \
0        False  {'variables': {'machineId': 'E5UW', 'orgArea':...
1        False  {'variables': {'machineId': 'E10UW', 'orgArea'...
2        False  {'variables': {'machineId': 'HAND4', 'orgArea'...
3         True  {'variables': {'machineId': '10165', 'orgArea'...
4        False  {'variables': {'machineId': 'E1UW', 'orgArea':...
...        ...                                                ...
11257    False  {'variables': {'machineId': '2683', 'roleId': ...
11258    False  {'variables': {'machineId': '2721', 'roleId': ...
11259    False  {'variables': {'machineId': '2748', 'roleId': ...
11260    False  {'variables': {'machineId': '2715', 'roleId': ...
11261    False  {'variables': {'machineId': '2709', 'roleId': ...

                                 title             headline  \
0                        E5UW broke down  Machine interrupted
1                       E10UW broke down  Machine interrupted
2                       HAND4 broke down  Machine interrupted
3                       10165 broke down  Machine interrupted
4                        E1UW broke down  Machine interrupted
...                                  ...                 ...
11257  2683   undefinierter Stillstand  Maschinenstillstand
11258  2721   undefinierter Stillstand  Maschinenstillstand
11259  2748   undefinierter Stillstand  Maschinenstillstand
11260  2715   undefinierter Stillstand  Maschinenstillstand
11261  2709   undefinierter Stillstand  Maschinenstillstand
```

```
                                      ruleId  autostart  \
0      {'$oid': '5d444be32eb5d80006d704fa'}       True
1      {'$oid': '5d444be32eb5d80006d704fa'}       True
2      {'$oid': '5d444be32eb5d80006d704fa'}       True
3      {'$oid': '5d444be32eb5d80006d704fa'}       True
4      {'$oid': '5d444be32eb5d80006d704fa'}       True
...                                     ...        ...
11257  {'$oid': '5d664c91fb35b900061d85c7'}       True
11258  {'$oid': '5d664c91fb35b900061d85c7'}       True
11259  {'$oid': '5d664c91fb35b900061d85c7'}       True
11260  {'$oid': '5d664c91fb35b900061d85c7'}       True
11261  {'$oid': '5d664c91fb35b900061d85c7'}       True

                                     project  \
0      {'$oid': '5d44361a5a5cbc00067a5f41'}
1      {'$oid': '5d44361a5a5cbc00067a5f41'}
2      {'$oid': '5d44361a5a5cbc00067a5f41'}
3      {'$oid': '5d44361a5a5cbc00067a5f41'}
4      {'$oid': '5d44361a5a5cbc00067a5f41'}
...                                     ...
11257  {'$oid': '5d44361a5a5cbc00067a5f41'}
11258  {'$oid': '5d44361a5a5cbc00067a5f41'}
11259  {'$oid': '5d44361a5a5cbc00067a5f41'}
11260  {'$oid': '5d44361a5a5cbc00067a5f41'}
11261  {'$oid': '5d44361a5a5cbc00067a5f41'}

                                lastNotification
0      {'$date': '2019-08-12T20:38:00.709+0000'}
1      {'$date': '2019-08-12T20:38:00.700+0000'}
2      {'$date': '2019-08-12T20:38:00.718+0000'}
3      {'$date': '2019-08-12T20:25:06.422+0000'}
4      {'$date': '2019-08-12T20:38:00.723+0000'}
...                                          ...
11257  {'$date': '2019-10-09T08:13:27.582+0000'}
11258  {'$date': '2019-10-09T07:57:06.704+0000'}
11259  {'$date': '2019-10-09T08:04:22.371+0000'}
11260  {'$date': '2019-10-09T08:09:01.134+0000'}
11261  {'$date': '2019-10-09T08:12:48.484+0000'}

[11262 rows x 13 columns]
```

[186]:
```
import matplotlib.pyplot as plt; plt.rcdefaults()
import numpy as np
import matplotlib.pyplot as plt

historic_tasks = data_mi_history._id.count()
```

```
tasks = data_mi._id.count()

objects = ('Historic MI Tasks', 'MI Tasks')
y_pos = np.arange(len(objects))
performance = [historic_tasks, tasks]

plt.bar(y_pos, performance, align='center', alpha=0.5)
plt.xticks(y_pos, objects)
plt.ylabel('Usage')
plt.title('Programming language usage')

plt.show()
```



[187]:
```
data_payload = pd.DataFrame(data_mi.payload.values.tolist())
data_vars  = pd.DataFrame(data_payload.variables.values.tolist())

data_date = pd.DataFrame(data_mi.lastNotification.values.tolist())
data_date = data_date.rename(columns={"$date": "date_new"})

data_id = pd.DataFrame(data_mi._id.values.tolist())
data_id = data_id.rename(columns={"$oid": "id_new"})
```

```
data_mi = data_mi.assign(machineId = data_vars.machineId)
data_mi = data_mi.assign(orgArea = data_vars.orgArea)
data_mi = data_mi.assign(interruptionId = data_vars.interruptionId)
data_mi = data_mi.assign(item = data_vars.item)
data_mi = data_mi.assign(site = data_vars.site)
data_mi = data_mi.assign(state = data_vars.state)
data_mi = data_mi.assign(status = data_vars.status)
data_mi = data_mi.assign(item = data_vars.item)
data_mi = data_mi.assign(counter = data_vars.counter)
data_mi = data_mi.assign(identifier = data_vars.identifier)
data_mi = data_mi.assign(id_new = data_id.id_new)
data_mi = data_mi.assign(date_ = data_date.date_new)

data_mi = data_mi.drop(['_id', 'isApp', 'source', 'deleted', 'headline',␣
 ↪'ruleId', 'autostart', 'project', 'lastNotification', 'payload',␣
 ↪'archived'], axis=1)

data_mi
```

[187]:

	status		title	machineId	orgArea
0	NaN		E5UW broke down	E5UW	QS
1	NaN		E10UW broke down	E10UW	QSS
2	NaN		HAND4 broke down	HAND4	ZUS_2
3	NaN		10165 broke down	10165	ZUS_2
4	NaN		E1UW broke down	E1UW	QS
...	...		...	...	...
11257	NaN	2683	undefinierter Stillstand	2683	NaN
11258	NaN	2721	undefinierter Stillstand	2721	NaN
11259	NaN	2748	undefinierter Stillstand	2748	NaN
11260	NaN	2715	undefinierter Stillstand	2715	NaN
11261	NaN	2709	undefinierter Stillstand	2709	NaN

	interruptionId	item	site	state	counter	identifier
0	null	F 99185	BRU	P_INTERRUPT	NaN	NaN
1	null	FP B3174	BRU	P_INTERRUPT	NaN	NaN
2	/	F B6027	BRU	P_INTERRUPT	NaN	NaN
3	null	F B5152	BRU	P_INTERRUPT	NaN	NaN
4	null	FV B9381	BRU	P_INTERRUPT	NaN	NaN
...	...	...	...	...	...	...
11257	NaN	NaN	NaN	PRODUCTION	NaN	23305
11258	NaN	NaN	NaN	PRODUCTION	NaN	21762
11259	NaN	NaN	NaN	PRODUCTION	NaN	16943
11260	NaN	NaN	NaN	PRODUCTION	NaN	16948
11261	NaN	NaN	NaN	PRODUCTION	NaN	302

	id_new	date_
0	5d4d80813a007a0006292b38	2019-08-12T20:38:00.709+0000

```
1       5d4d80813a007a0006292b3a  2019-08-12T20:38:00.700+0000
2       5d4d80813a007a0006292b3d  2019-08-12T20:38:00.718+0000
3       5d4d80813a007a0006292b40  2019-08-12T20:25:06.422+0000
4       5d4d80813a007a0006292b44  2019-08-12T20:38:00.723+0000
...                          ...                           ...
11257   5d9d92b4e59c620007b76737  2019-10-09T08:13:27.582+0000
11258   5d9d92d2e59c620007b7673e  2019-10-09T07:57:06.704+0000
11259   5d9d92fbe59c620007b76749  2019-10-09T08:04:22.371+0000
11260   5d9d9318e59c620007b7674e  2019-10-09T08:09:01.134+0000
11261   5d9d9680e59c620007b768dd  2019-10-09T08:12:48.484+0000

[11262 rows x 12 columns]
```

1.1 Visualisation

[188]:
```
data_machineId = data_mi.machineId.value_counts()
data_machineId.iplot(kind='bar',title='Per Machine Id')
```

[189]:
```
data_date = data_mi.date_.value_counts()
data_date.iplot(kind='bar',title='Per day')
```

[190]:
```
data_interruption = data_mi.interruptionId.value_counts()
data_interruption.iplot(kind='bar',title='per interruption')
```

[191]:
```
data_state = data_mi.state.value_counts()
data_state.iplot(kind='bar',title='per state')
```

[193]:
```
data_orgarea = data_mi.orgArea.value_counts()
data_orgarea.iplot(kind='bar',title='per orgArea')
```