Facebook Check-In Project

Contents

Introduction

Companies are increasingly looking to social media platforms like Facebook to collect data about (potential) customers' attitudes and engagement. For example, Facebook has a feature that allows users to "check in" to various establishment/business locations. Data about these check-ins are available for analysis.

Check-ins are (arguably) not just a proxy for counting the number of people who visit establishments, they are also a form of advertising to the customer's friends on Facebook. Therefore this metric can be a strong performance indicator for the business and may help in making performance predictions (e.g. stock prices).

In this project we will analyze a dataset from Thinknum about Facebook check-ins.

Exploratory Data Analysis

The dataset is available at https://blog.thedataincubator.com/tag/data-sources/ in the "Facebook Followers" section. In this section we import the data and do a very preliminary analysis.

```
% imported csv as 'fb'
size(fb)
```

```
ans = 1×2
    3621391          11
```
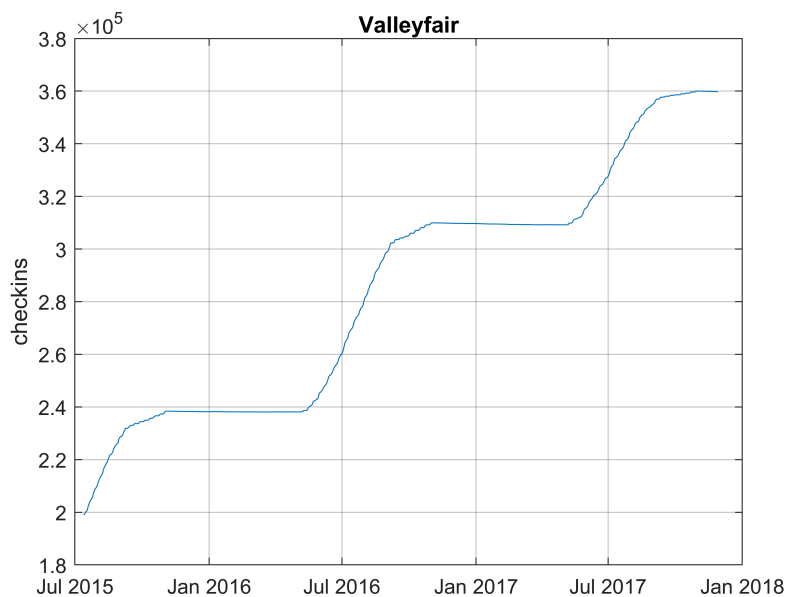
Let us look at two examples:

```
cindex = fb.username == 'Disneyland';
plot(fb.time(cindex),fb.checkins(cindex))
title('Disneyland')
ylabel('checkins')
```
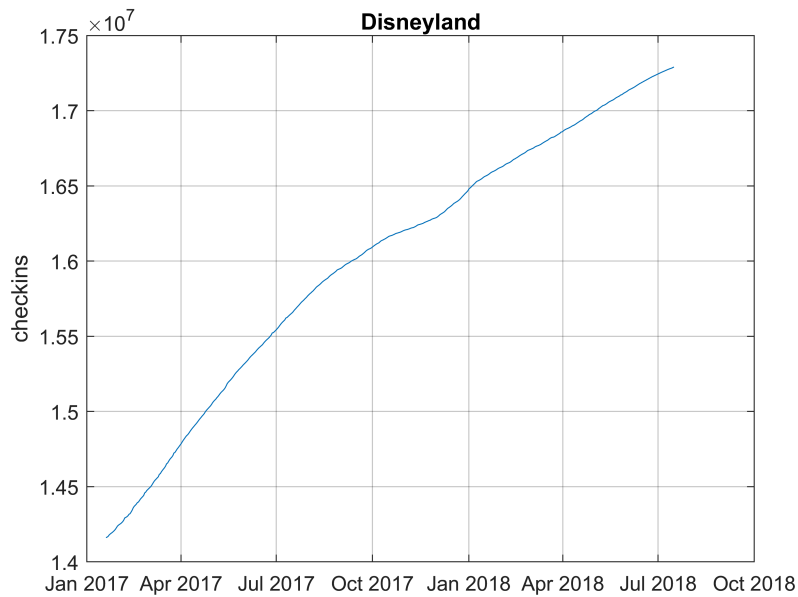
The Valleyfair amusement park seems to be closed in the winter:

```
cindex = fb.username == 'valleyfair';
plot(fb.time(cindex),fb.checkins(cindex))
title('Valleyfair')
ylabel('checkins')
```


Which businesses get the most check-ins?

```
undefined = isundefined(fb.username);
fb_def = fb(~undefined,:);
companies = unique(fb_def.username);
checkins_c = [];
```

```
for c = companies'
    cindex = fb.username == c;
    checkins_c(end+1) = max(fb.checkins(cindex));
end
checkins_c = checkins_c';
```

```
company_max_checkins = table(companies,checkins_c);
most = sortrows(company_max_checkins,2,"descend");
top10 = most(1:10,:)
```

top10 = 10×2 table

	companies	checkins_c
1	Disneyland	17290550
2	WaltDisneyW...	6547775
3	DisneylandP...	4487255
4	disneylandp...	4063748
5	TheGarden	3651157
6	MadisonSqua...	3282939
7	empirestate...	2960168
8	KnottsBerry...	2947001
9	SeaWorld	2916920
10	BuschGarden...	2745204

```
labels = top10.companies;
temp = cellstr(labels);
ctemp = categorical(temp);
labels = reordercats(ctemp,temp);
```

```
bar(labels,top10.checkins_c)
title('Companies with the most checkins (top 10)')
ylabel('checkins')
```

Companies with the most checkins (top 10)