Project 3 Proposal: classification of start-up companies

Business Case:
There are many start-up companies founded every year. There are start-ups, which are acquired by other companies, while others remain independent or decide to be publicly trade.

There are economic implications for any of these outcomes for the investors, employees and founders. Thus, developing a model to classify start-ups would be helpful to these stakeholders.

Minimum Viable Product:

The MVP for this project is to classify companies five years and older between acquired and not acquired.

The Data:

The data is a 2013 snapshot contains a replica of the Crunchbase dataset from October 2013. That includes all organization, and product profiles along with funding rounds, acquisitions, and IPOs.

crunchbase-companies.csv	Variable used	Alternative variable to be created
name	Y	
category_code	Y	
funding_total_usd	N	
status	Y	
country_code	Y	
state_code	Y	
region	Y	
city	Y	
funding_rounds	Y	
founded_at	Y	Estimated to calculate other variables and filter into companies less than 5 years old
founded_month	Y	
founded_quarter	Y	
founded_year	N	
first_funding_at	Y	
last_funding_at	Y	Last_funding (in days since company foundation)
last_milestone_at	N	

crunchbase-acquisitions.csv		
company_name	Y	
company_category_code	Y	
company_country_code	Y	
company_state_code	Y	
company_region	Y	
company_city	Y	
acquirer_permalink	N	
acquirer_name	N	
acquirer_category_code	N	
acquirer_country_code	N	
acquirer_state_code	N	
price_currency_code	N	
acquirer_region	N	
acquirer_city	N	
acquired_at	N	
acquired_month	N	
acquired_quarter	N	
acquired_year	N	
price_amount	N	
crunchbase-investments		
company_name	Y	
company_category_code	Y	
company_country_code	Y	
company_state_code	Y	
company_region	Y	
company_city	Y	
investor_permalink	N	
investor_name	N	
investor_category_code	N	
investor_country_code	N	
investor_state_code	N	
investor_region	N	
investor_city	N	
funding_round_type	Y	
funded_at	Y	Relative to founding date
funded_month	N	
funded_quarter	N	
funded_year	N	

`raised_amount_usd`	Y	
crunchbase-rounds.csv		
`company_name`	Y	
`company_category_code`	Y	
`company_country_code`	Y	
`company_state_code`	Y	
`company_region`	Y	
`company_city`	Y	
`funding_round_type`	Y	
`funded_at`	Y	Relative to founding date
`funded_month`	N	
`funded_quarter`	N	
`funded_year`	N	
raised_amount_usd	Y	

crunchbase-companies.csv
```
name                17727 non-null object
category_code       17328 non-null object
funding_total_usd   16222 non-null object
status              17727 non-null object
country_code        17727 non-null object
state_code          17480 non-null object
region              17727 non-null object
city                17462 non-null object
funding_rounds      17727 non-null int64
founded_at          13676 non-null object
founded_month       13676 non-null object
founded_quarter     13676 non-null object
founded_year        13676 non-null float64
first_funding_at    17622 non-null object (time after foundation)
last_funding_at     17622 non-null object
last_milestone_at   10528 non-null object
```

crunchbase-acquisitions.csv
```
company_name            4651 non-null object
company_category_code   3933 non-null object
company_country_code    4651 non-null object
company_state_code      4543 non-null object
company_region          4651 non-null object
company_city            4532 non-null object
acquirer_permalink      4651 non-null object
acquirer_name           4651 non-null object
acquirer_category_code  4137 non-null object
acquirer_country_code   4384 non-null object
acquirer_state_code     3877 non-null object
acquirer_region         4651 non-null object
acquirer_city           4262 non-null object
```

```
acquired_at                4651 non-null object
acquired_month             4651 non-null object
acquired_quarter           4651 non-null object
acquired_year              4651 non-null float64
price_amount               1409 non-null float64
price_currency_code        4648 non-null object
```

crunchbase-investments
```
company_name               52869 non-null object
company_category_code      52227 non-null object
company_country_code       52869 non-null object
company_state_code         52378 non-null object
company_region             52869 non-null object
company_city               52337 non-null object
investor_permalink         52868 non-null object
investor_name              52868 non-null object
investor_category_code      2443 non-null object
investor_country_code      40869 non-null object
investor_state_code        36061 non-null object
investor_region            52868 non-null object
investor_city              40390 non-null object
funding_round_type         52867 non-null object
funded_at                  52867 non-null object
funded_month               52867 non-null object
funded_quarter             52867 non-null object
funded_year                52867 non-null float64
raised_amount_usd          49271 non-null float64
```

crunchbase-rounds.csv
```
company_name               31679 non-null object
company_category_code      31215 non-null object
company_country_code       31679 non-null object
company_state_code         31334 non-null object
company_region             31679 non-null object
company_city               31329 non-null object
funding_round_type         31679 non-null object
funded_at                  31679 non-null object
funded_month               31679 non-null object
funded_quarter             31679 non-null object
funded_year                31679 non-null int64
raised_amount_usd          29100 non-null float64
```