Adriano Pais Rodrigues m20180430
(Long list answers were formatted by hand)
1. Preliminary Analysis
    a. Check the code
    b. Questions
        i.   Action,Adventure,Sci-Fi at **38236427** votes total
        ii. (Music,News,Short) and (Biography,History,Mystery) **tie** at **9.1**
        iii. Most common genre is **Drama** at **1548377**

        iv. 2017, Drama, 73985
             2015, Comedy, 70808
             2025, Action, 2
             2014, Comedy, 69185
             2012, Drama, 63530
             2008, Drama, 48956
             2115, Drama, 1
             2018, Drama, 65707
             2013, Comedy, 66357
             2003, Drama, 30587
             2020, Drama, 393
             2019, Drama, 22227
             2007, Drama, 46722
             2022, Action, 8
             2023, Action, 1
             2001, Drama, 25138
             2005, Drama, 36197
             2002, Drama, 26119
             2004, Drama, 30547
             2021, Animation, 24
             2016, Comedy, 73728
             2010, Drama, 55342
             2011, Drama, 59558
             2006, Drama, 41214
             2009, Drama, 50680
             2024, Fantasy, 2
             2000, Drama, 22204
        v.
2. Network inference
    a. -
        i.   _2_a_i_df on Cmd 18
        ii. _2_a_ii_df on Cmd 19
        iii. there are **82505** actors and actresses
        iv. _2_a_iv_df in the code
    b.
        i.   -
            1. Top10:

```
---+------------------+
|       id|         pagerank|
+---------+------------------+
|nm0000616|  36.7585650511952|
|nm0000514|21.774708486556175|
|nm0001744|20.781573222871835|
|nm0001803|19.045207082969615|
|nm0001698| 16.94617108138148|
|nm0000448|15.969963708601718|
|nm0004193|15.551239274891962|
|nm0000367| 15.52797225280788|
|nm0626259|15.514587903780885|
|nm0261724|15.346684753639488|
```

```
                  +---------+------------------+
     2. Botom 10: +-------------------+
                  |       id|          pagerank|
                  +---------+------------------+
                  |nm0762643|0.17402301977110346|
                  |nm7987433| 0.1754137142786428|
                  |nm0080295|0.17613201131708234|
                  |nm0008711|0.17622469923828654|
                  |nm0012868|0.17622469923828654|
                  |nm3576158|0.17634048023943583|
                  |nm0451809|0.17634048023943583|
                  |nm6000269| 0.1775649905701078|
                  |nm0481568| 0.1775649905701078|
                  |nm2056531| 0.1781407466057339|
                  +---------+------------------+
```

ii.   y: 1953, pr: 1.177   y: 2002, pr: 0.638

y: 1957, pr: 1.161   y: 2009, pr: 0.423

y: 1987, pr: 0.806   y: 1995, pr: 0.664

y: 1956, pr: 1.143   y: 1964, pr: 1.108

y: 1936, pr: 1.124   y: 1946, pr: 1.212

y: 2012, pr: 0.293   y: 2006, pr: 0.567

y: 1958, pr: 1.130   y: 1927, pr: 0.906

y: 1943, pr: 1.117   y: 1921, pr: 0.660

y: 1972, pr: 1.024   y: 1976, pr: 0.949

y: 1931, pr: 1.076   y: 1942, pr: 1.085

y: 1926, pr: 0.910   y: 1935, pr: 1.129

y: 1988, pr: 0.770   y: 1947, pr: 1.150

y: 1938, pr: 1.134   y: 1968, pr: 1.050

y: 1932, pr: 0.921   y: 1967, pr: 1.051

y: 1977, pr: 0.889   y: 1922, pr: 0.664

y: 1971, pr: 1.001   y: 1924, pr: 0.856

y: 2014, pr: 0.526   y: 2004, pr: 0.570

y: 1984, pr: 0.849   y: 1939, pr: 1.182

y: 1982, pr: 0.880   y: 1989, pr: 0.779

y: 1941, pr: 1.201   y: 2011, pr: 0.618

y: 2013, pr: 1.319   y: 1992, pr: 0.740

y: 2005, pr: 0.553   y: 1961, pr: 1.112

y: 2000, pr: 0.582   y: 1951, pr: 1.144

y: 1965, pr: 1.112   y: 1966, pr: 1.086

y: 1962, pr: 1.079   y: 2008, pr: 0.440

y: 1954, pr: 1.173   y: 1955, pr: 1.242

y: 1930, pr: 1.138   y: 1999, pr: 0.592

y: 1981, pr: 0.906   y: 1963, pr: 1.124

y: 1940, pr: 1.152   y: 1994, pr: 0.682

y: 1978, pr: 0.929   y: 1950, pr: 1.148

y: 1974, pr: 0.956   y: 1997, pr: 0.668

y: 1959, pr: 1.124   y: 1973, pr: 0.983

```
                  y: 1925, pr: 0.915                    y: 1979, pr: 0.899
                  y: 2007, pr: 0.479                    y: 2015, pr: 0.668
                  y: 1996, pr: 0.701                    y: 1998, pr: 0.641
                  y: 1983, pr: 0.862                    y: 1949, pr: 1.120
                  y: 1969, pr: 1.027                    y: 1993, pr: 0.708
                  y: 1923, pr: 0.943                    y: 2001, pr: 0.602
                  y: 1980, pr: 0.895                    y: 2010, pr: 0.504
                  y: 1944, pr: 1.096                    y: 1990, pr: 0.752
                  y: 1960, pr: 1.060                    y: 1928, pr: 0.901
                  y: 1934, pr: 1.017                    y: 1991, pr: 0.735
                  y: 1937, pr: 1.127                    y: 2003, pr: 0.581
                  y: 1948, pr: 1.165                    y: 1945, pr: 1.133
                  y: 1986, pr: 0.813                    y: 1975, pr: 0.922
                  y: 1952, pr: 1.137                    y: 1970, pr: 0.986
                  y: 1985, pr: 0.808                    y: 1933, pr: 1.049
                  y: 1929, pr: 1.121
```

c. -

    i.   Top 10:

```
+------+---------+------------------+
|degree|       id|          pagerank|
+------+---------+------------------+
|   852|nm0000616|  36.7585650511952|
|   440|nm0001698| 16.94617108138148|
|   238|nm0004193|15.551239274891962|
|   270|nm0261724|15.346684753639488|
|   550|nm0000514|21.774708486556175|
|   526|nm0000367| 15.52797225280788|
|   400|nm0626259|15.514587903780885|
|   456|nm0001803|19.045207082969615|
|   392|nm0000448|15.969963708601718|
|   500|nm0001744|20.781573222871835|
+------+---------+------------------+
```

    ii.   Bottom 10:

```
------+------------------+
|degree|       id|          pagerank|
+------+---------+------------------+
|     2|nm0012868|0.17622469923828654|
|     2|nm3576158|0.17634048023943583|
|     2|nm0762643|0.17402301977110346|
|     2|nm6000269| 0.1775649905701078|
|     2|nm7987433| 0.1754137142786428|
|     2|nm0008711|0.17622469923828654|
|     2|nm0451809|0.17634048023943583|
|     2|nm0080295|0.17613201131708234|
|     2|nm2056531| 0.1781407466057339|
|     2|nm0481568| 0.1775649905701078|
+------+---------+------------------+
```

3. The Kevin Bacon Game

    a. -

        i.   nm0000102 is Kevin Bacon's nconst (id)

      ii.   gf.shortestPaths(landmarks=['nm0000102']) # Cmd 33

    iii.

```
+-----+-----+
|value|count|
+-----+-----+
|    1|  126|
|    6| 2097|
|    3|17817|
|    5|12652|
|    9|   18|
|    4|28565|
|    8|   45|
|    7|  300|
|   10|    4|
|   11|    2|
|    2| 3258|
|    0|    1| <- That's Kevin Bacon!
+-----+-----+
```

Also, 6204 actors not connected to Kevin

4. Engineering the perfect cast
   a. -
      i. That is either df_title_ratings from the beginning of the code or I didn't understand the question.
     ii. _4_a_ii on Cmd 36
    iii. _4_a_iii on Cmd 37
     iv. Check Cmd 42
      v. **There is no v question**
     vi. Model intercept: 6.0353589330473945, coefficient: -0.082826843805465164