

Artists deaths and their impact on artwork sales

INTRODUCTION

Rumour says, that CD stores had “Johnny Halliday : The Collector edition” kept in stock, in the back, prepared for the day Johnny would pass, so that they would be prepared for the massive demand. This phenomenon – A sudden peak of interest in one’s artwork after his death – is described in a number of articles on the web. Can we statistically highlight this effect using data science ?

We will use amazon reviews data (products reviews from 2003 to 2013) utilizing the number of reviews over a time period as an interest metric, artists lifespan data scraped from Wikipedia, and amazon products metadata dataset to link the previous two. The results are analyzed using time series analysis techniques.

DATA GATHERING

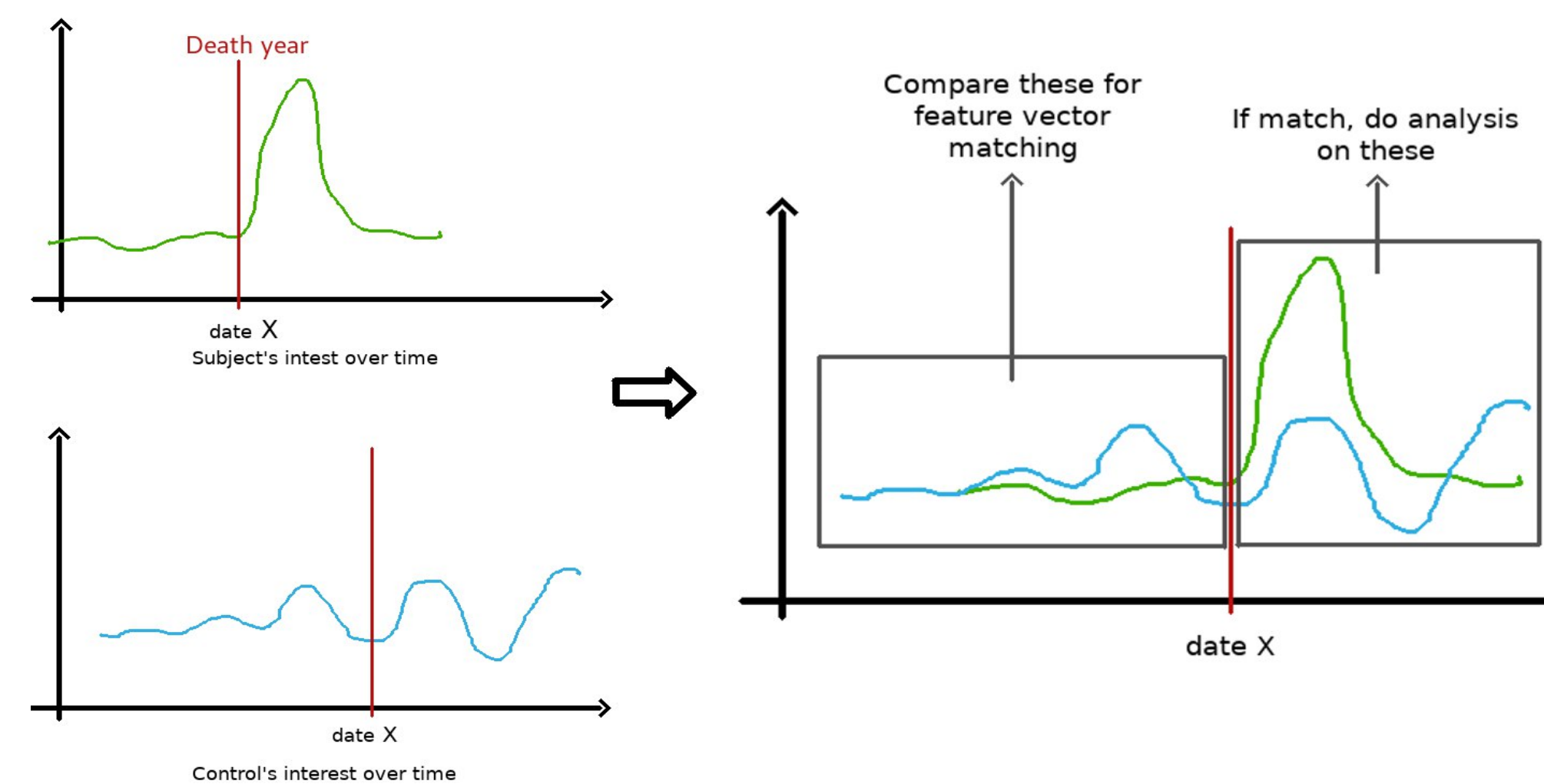
Amazon meta data was completed with Amazon API queries to get the artists’ names corresponding to a specific product ID. Then, combining this information with Amazon review data, It was possible to associate artists’ names with reviews count.

The list of dead artists was scrapped from Wikipedia. It allowed separating the subjects, which were dead in the required time period, from the controls. Finally, deriving reviews count in time from this data was then relatively trivial.

DATA ANALYSIS

To correct for the general increase in amazon usage over time, each time series was normalized by a general reviews count tendency curve.

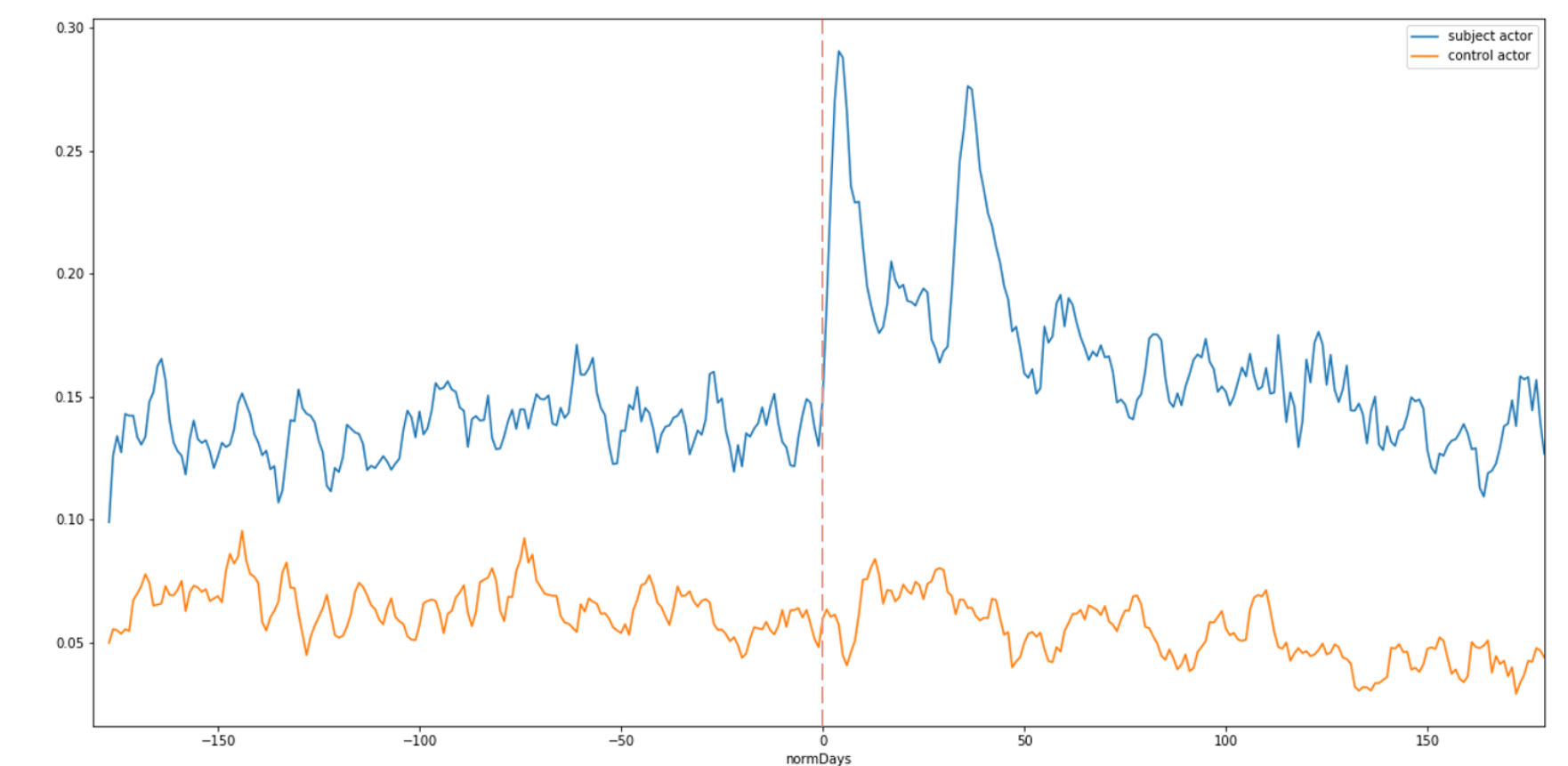
Then, pairs of subject and control were constructed so that the two artists have a similar time series, from 6 months to the death date of the subject. Then, each subject time series was truncated from 6 months before to 6 months after its death date and the same window was applied to its matched control.



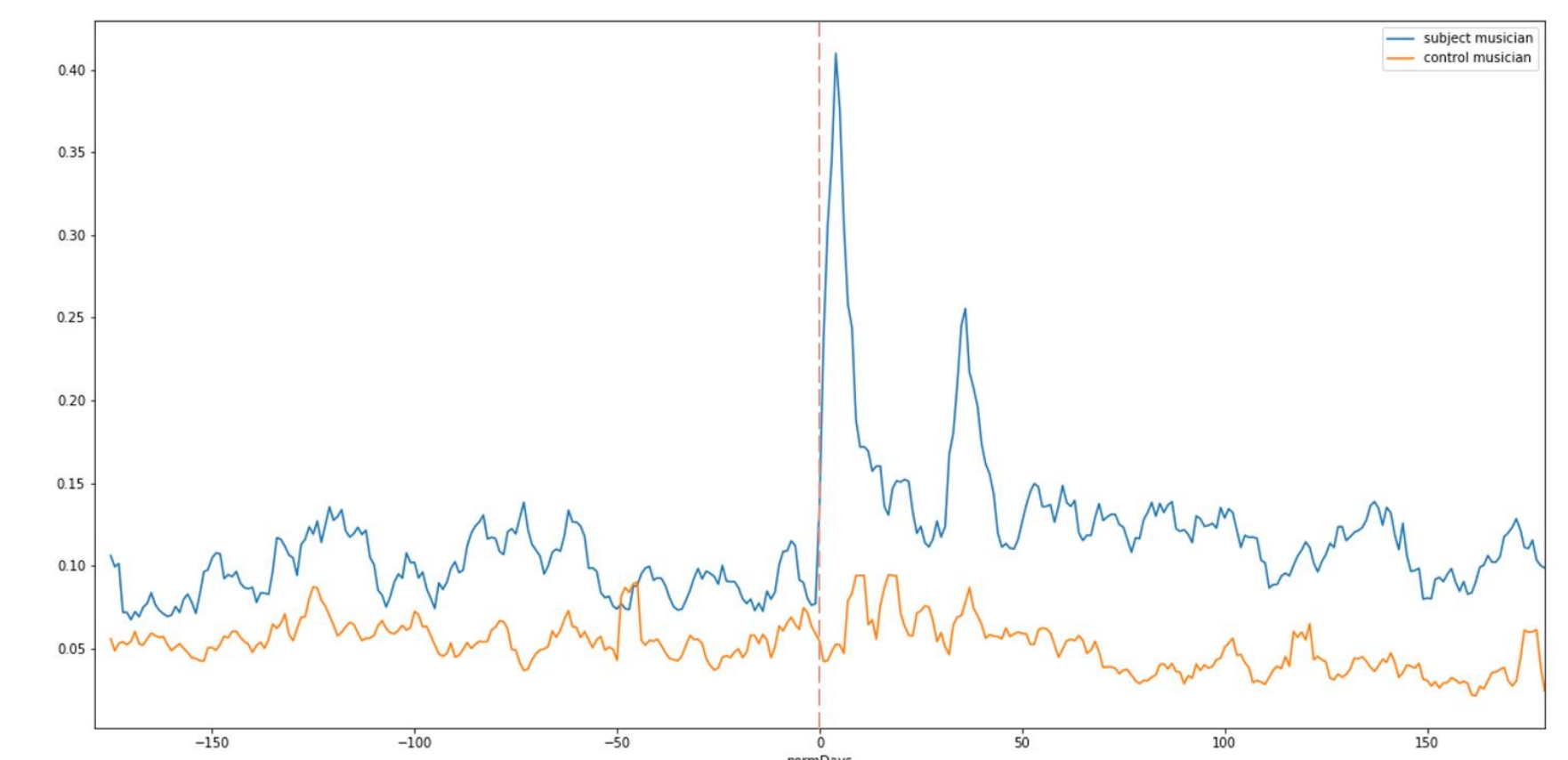
Controls and subjects are matched on interest over time, centered at the artist's death date

RESULTS

Reviews frequency for actors

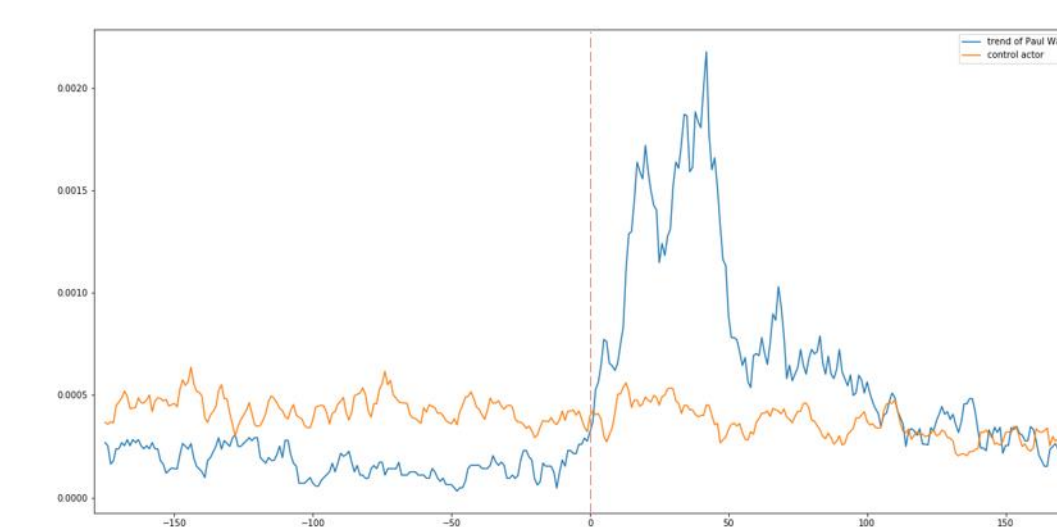


Reviews frequency for musicians



BONUS RESULTS

Paul Walker



Whitney Houston

