



UNIVERSIDAD
DE LA REPÚBLICA
URUGUAY

La obra de William Shakespeare: un breve análisis

Tarea 1 Introducción a la ciencia de datos

Alvaro Valiño - Lucía Coudet

Montevideo, Mayo 2024

1. Introducción

William Shakespeare fue un dramaturgo, poeta, y actor de origen inglés nacido en abril del año 1564 en *Warwickshire* al sureste de *Birmingham*, Inglaterra.

Cursó sus primeros estudios en la escuela primaria local *Stratford Grammar School* lo cual se piensa que le aportó una educación sólida en gramática y literatura. A pesar de que la calidad de las escuelas gramaticales de la época era bastante irregular, existen indicios en el sentido de que la de *Stratford* era bastante buena.

El 28 de noviembre de 1582, cuando tenía dieciocho años de edad, contrajo matrimonio con Anne Hathaway, de 26 años y originaria de una localidad próxima a *Stratford*.

Alrededor del año 1592 comenzó su carrera teatral en Londres escribiendo y actuando en obras para la compañía teatral *Lord Chamberlain's Men* la cual fue mas tarde conocida como *King's Men*.

Sus obras que incluyen tragedias como *Hamlet* y *Macbeth*, comedias como *El sueño de una noche de verano* y *La comedia de los errores* y obras históricas como *Enrique V* y *Ricardo III*, entre otras, son celebradas por su profundidad, ingenio, y habilidad para explorar la complejidad humana.

William Shakespeare se convirtió en una figura central de la literatura inglesa y su legado continua siendo influyente en la actualidad.

2. Objetivo

Este trabajo tiene por objetivo presentar un breve análisis de la obra Shakespeariana utilizando las bases de datos disponibles en <https://relational-data.org/dataset/Shakespeare>.

Se trata de una base de datos relacional que cuenta con información sobre las obras, los capítulos, los personajes y los párrafos para cada personaje de cada obra.

3. Análisis exploratorio de datos

William Shakespeare produjo un total de 43 obras entre los años 1589 y 1612 y las mismas abarcan los géneros de tragedia, comedia, historia, poesía y soneto.

En el cuadro 1 del anexo se presentan la totalidad de sus obras en orden cronológico según la fecha de realización de las mismas. Como puede observarse en el mismo, las obras varían entre los distintos géneros a través de los años.

En la Figura 1 se presenta el gráfico de barras de los géneros de sus obras. En el mismo puede observarse los 3 géneros predominantes entre sus obras son comedia, historia y tragedia.

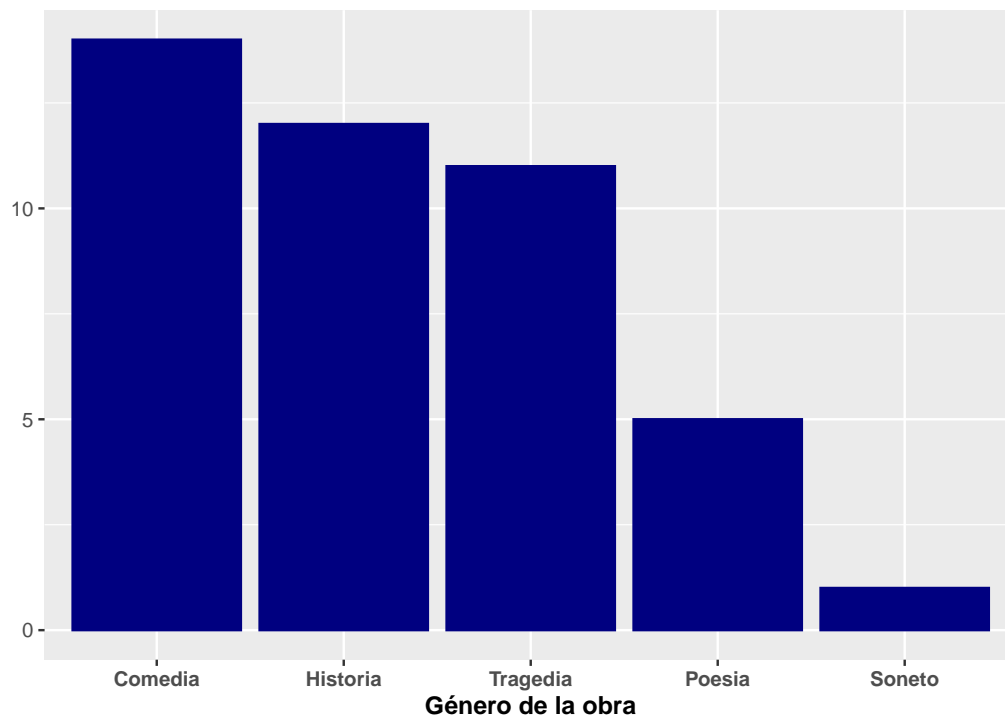


Figura 1: Gráfico de barras de los géneros de las obras de William Shakespeare

Por otra parte, en lo que respecta a la composición a lo largo de los años, el gráfico de barras de la Figura 2 sugiere que no existe una tendencia clara para un género en particular. Asimismo, se observa que en todos los años entre 1589 y 1612 produjo obras escritas excepto en el año 1603 donde no existe registro de la producción de ninguna obra.

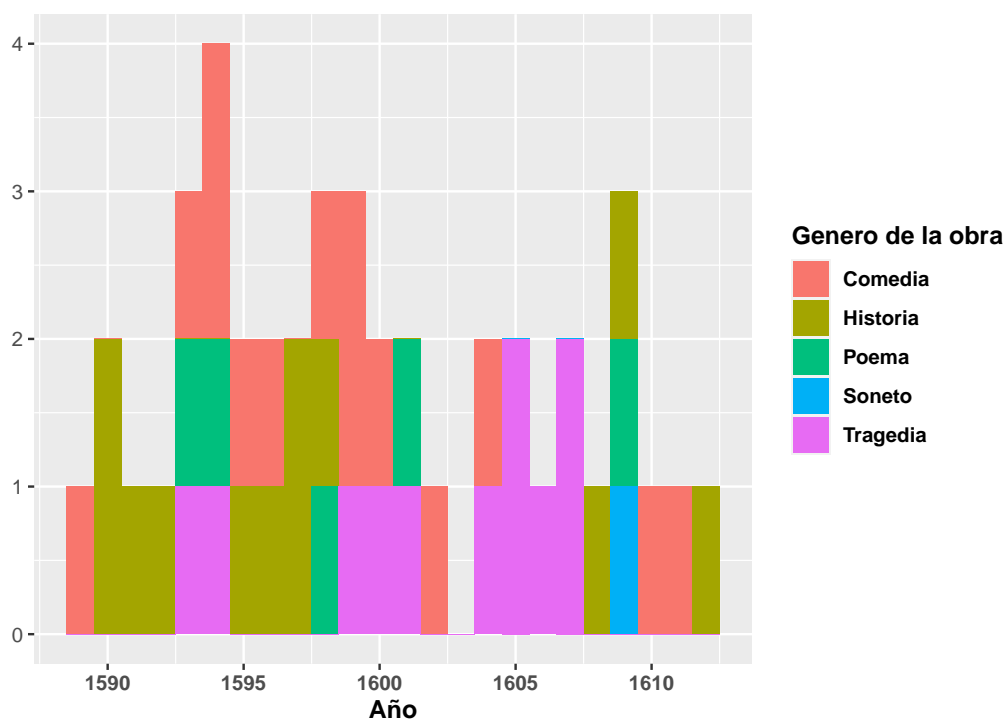


Figura 2: Gráfico de barras de las obras de Shakespeare a través del tiempo, según el género de las mismas

4. Algunas de sus obras mas relevantes

Hamlet y *Romeo y Julieta* son dos de las obras mas reconocidas de Shakespeare. *Romeo y Julieta* (1594) es una tragedia donde se cuenta la historia de dos jóvenes amantes, *Romeo* y *Julieta*, cuyas familias estan enemistadas. A pesar de las rivalidades de sus familias, se enamoran profundamente y deciden casarse en secreto. Esta obra explora temas como el amor, el destino, la violencia, y el conflicto familiar.

Por otro lado, *Hamlet* (1600) es también una tragedia que gira en torno al Príncipe Hamlet de Dinamarca, quien busca vengar la muerte de su padre, el Rey, en manos de su tío Claudio. Luego de la muerte del Rey, su hermano Claudio se convierte en el nuevo Rey y se casa con su esposa, la madre de *Hamlet*. Este obra examina temas como la venganza, la locura, la traicion y la corrupción.

Como fue mencionado anteriormente, la base de datos de los *párrafos* contiene la informacion sobre lo que dice cada personaje en cada acto de cada escena. Si analizamos la cantidad de veces que interviene cada personaje en la obra de *Romeo y Julieta*, vemos como el personaje con mas intervenciones es *Romeo* seguido por *stage directions*. Es recién en tercer lugar que se encuentra *Julieta*. En total, estos tres personajes tienen el 43.2% de las intervenciones de la obra.

Cuadro 1: Intervenciones de cada personaje en la obra Romeo y Julieta

CharName	conteo	porc
Romeo	163	16.5
(stage directions)	146	14.8
Juliet	118	11.9
Nurse	90	9.1
Benvolio	64	6.5
Mercutio	62	6.3
Friar Laurence	55	5.6
Capulet	51	5.2
Lady Capulet	45	4.6
Paris	23	2.3

Por otro lado, en lo que respecta a la obra de *Hamlet*, el personaje Hamlet tiene el 28,1 % de las intervenciones seguido por *stage directions* que tiene el 11 %. En total, estos dos personajes tienen en total el 39.1 % de las intervenciones.

Cuadro 2: Intervenciones de cada personaje en la obra Hamlet

CharName	conteo	porc
Hamlet	358	28.1
(stage directions)	140	11.0
Horatio	109	8.5
Claudius	106	8.3
Polonius	86	6.7
Gertrude	72	5.6
Laertes	62	4.9
Ophelia	58	4.5
Rosencrantz	48	3.8
Marcellus	37	2.9

Como vemos tanto en la obra de *Romeo y Julieta* como en la obra de *Hamlet*, *stage directions* es el segundo personaje con mas intervenciones. Esto implica que el segundo personaje con mas intervenciones en ambas obras no es en si mismo un personaje, sino que son las acotaciones durante la obra. Esto es un aspecto que caracteriza a las obras de William Shakespeare.

Ahora bien, los resultados anteriores se obtuvieron contando el número de intervenciones y no el número de palabras. Una pregunta que se desprende del análisis anterior es cuales son los personajes con mayor número de palabras en las obras de Shakespeare.

Para ello, trabajamos con las bases de datos de los párrafos y de los personajes. En el gráfico de barras de la Figura 3 se presenta el % de palabras dichas por los 10 personajes con más palabras.

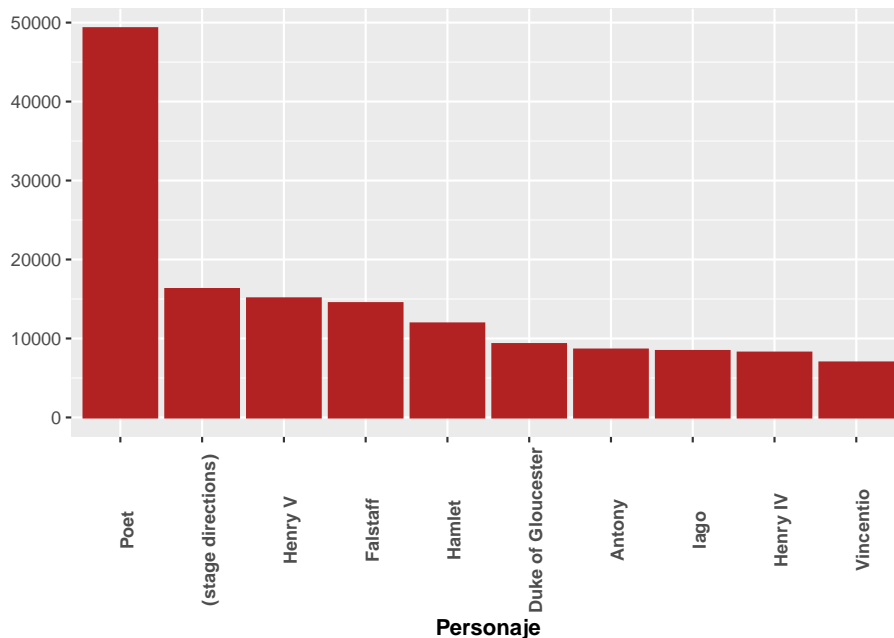


Figura 3: Los 10 personajes de las obras de Shakespeare con mas palabras

El gráfico nos muestra que los dos personajes con mas palabras son *Poet* y *stage directions* los cuales no son personajes en sí mismos. *Poet* es la voz de la poesia de Shakespeare mientras que, como fue mencionado anteriormente, *stage directions* son las acotaciones en el escenario.

Si excluimos a *Poet* y *stage directions* del análisis, en el gráfico de barras de la Figura 4 se presentan los 10 personajes con mas palabras.

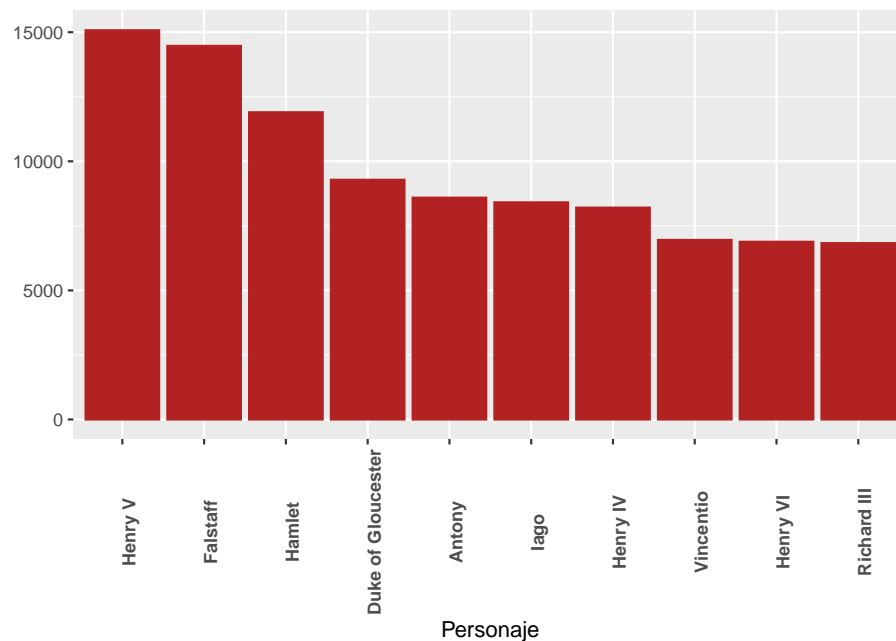


Figura 4: Los 10 personajes de las obras de Shakespeare con mas palabras excluyendo a Poet y Stage directions

Por último, en la figura 5 se presenta el gráfico de palabras de todas las obras de Shakespeare. Se observa que las palabras que mas aparecen son conectoras como por ejemplo *and*, *the*, *of*, entre otras.

Por lo tanto, para poder realizar un análisis de las palabras resulta necesario avanzar un paso más en la limpieza de la base de datos y eliminar las palabras que sean conectoras.

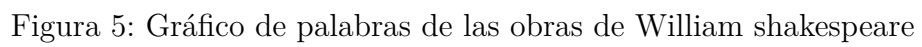
5. Comentarios finales

Las obras de William Shakespeare fueron producidas entre los años 1589 y 1612 siendo los géneros predominantes en sus obras comedia, historia, y tragedia.

Si se analiza a través de los años, no se observa alguna tendencia en particular en la selección de los géneros que produjo.

Si se analizan todas las obras en su conjunto, *Poet y stage directions* aparecen como los personajes con más palabras, los cuales no son en sí mismos personajes. Por otra parte, las palabras que más aparecen son conectoras lo cual es de esperar ya que no se excluyeron del análisis antes de realizar el gráfico de palabras.

Para futuros análisis se sugiera avanzar en la limpieza de los datos, en particular de las palabras.



6. Anexo

Cuadro 3: Las obras de William Shakespeare

Titulo	Fecha	Genero
Comedy of Errors	1589	Comedy
Henry VI, Part II	1590	History
Henry VI, Part III	1590	History
Henry VI, Part I	1591	History
Richard III	1592	History
Taming of the Shrew	1593	Comedy
Titus Andronicus	1593	Tragedy
Venus and Adonis	1593	Poem
Love's Labour's Lost	1594	Comedy
Rape of Lucrece	1594	Poem
Romeo and Juliet	1594	Tragedy
Two Gentlemen of Verona	1594	Comedy
Midsummer Night's Dream	1595	Comedy
Richard II	1595	History
King John	1596	History
Merchant of Venice	1596	Comedy
Henry IV, Part I	1597	History
Henry IV, Part II	1597	History
Henry V	1598	History
Much Ado about Nothing	1598	Comedy
Passionate Pilgrim	1598	Poem
Twelfth Night	1599	Comedy
As You Like It	1599	Comedy
Julius Caesar	1599	Tragedy
Hamlet	1600	Tragedy
Merry Wives of Windsor	1600	Comedy
Phoenix and the Turtle	1601	Poem
Troilus and Cressida	1601	Tragedy
All's Well That Ends Well	1602	Comedy
Measure for Measure	1604	Comedy
Othello	1604	Tragedy
King Lear	1605	Tragedy
Macbeth	1605	Tragedy
Antony and Cleopatra	1606	Tragedy
Coriolanus	1607	Tragedy
Timon of Athens	1607	Tragedy
Pericles	1608	History
Cymbeline	1609	History
Lover's Complaint	1609	Poem
Sonnets	1609	Sonnet
The Winter's Tale	1610	Comedy
Tempest	1611	Comedy
Henry VIII	1612	History