Food Recommendation system by using Principal component Analysis

[1] CH.Phani kumar [2]SD.Munwar Ali,[3] J.Vyuhitha ,[4] P. Vaishnavi,

[5]M. Mahesh raju

[1]*Asst. Professor, Dept. of CSE, Usha Rama College of Engineering and Technology, Telaprolu, Ap, India.*

[2,3,4,5] *dept. of CSE, Usha Rama College of Engineering and Technology, Telaprolu, Ap, India.*

Abstract

Food recommendation systems have gained immense popularity in recent years due to the convenience they offer in helping users discover and order their preferred dishes from a vast array of restaurants. In this research paper, we propose an innovative approach to enhance the efficiency and accuracy of food recommendations by integrating Principal Component Analysis (PCA) for dimensionality reduction, using Swiggy's extensive dataset as a case study.

The primary objective of this study is to improve user experience and increase the relevance of food recommendations. We achieve this by reducing the dimensionality of the dataset, making it more manageable and computationally efficient, while preserving the essential features for generating personalized recommendations. The Swiggy dataset used in this research encompasses a wide range of information, including restaurant ratings, city locations, menu item costs, restaurant names, cuisine types, and links to place orders.

Keyword's

Principal Component (PC), Dimensionality Reduction ,Variance ,eigen values ,eigen vectors ,covariance matrix

1. Introduction

The rapid growth of the food delivery industry has transformed the way people dine and order food. In this era of convenience, online food delivery platforms have become an integral part of urban life. Among these platforms, Swiggy stands out as a prominent player, providing access to an extensive network of restaurants and a diverse range of cuisines. As the competition in this domain intensifies, the need for more sophisticated and personalized food recommendation systems becomes increasingly evident.

Table 1: Attribute Information

Attribute	Type	Description
I'd	Discrete	Each food item have unique name
Name	Discrete	Names of the restaurants
City	Discrete	Name of the city
Cost	Continuous	Cost will be changed
Rating	Continuous	Rating is from 1-10
Cuisine	Continuous	Different type of food
Address	Discrete	Address of the restaurants
Link	Discrete	Link of restaurants in swiggy to order

Traditional food recommendation systems often rely on collaborative filtering, content-based filtering, or hybrid approaches to suggest dishes or restaurants to users. While these methods have shown promise, they are not without limitations. Collaborative filtering depends heavily on user interaction data and may struggle to provide meaningful recommendations for new users with limited history. Content-based filtering, on the other hand, tends to rely on extensive feature engineering, which can be labor-intensive and may not capture the nuanced preferences of users.

In response to these challenges, this research paper introduces an innovative approach to enhance the food recommendation system's efficiency and accuracy by incorporating Principal Component Analysis (PCA) for dimensionality reduction. PCA is a widely used technique in machine learning and data analysis for identifying the most informative features while reducing the computational complexity of large datasets.

The dataset employed in this study is sourced from Swiggy, one of India's leading food delivery platforms. It comprises a wealth of information, including restaurant ratings, city locations, menu item costs, restaurant names, cuisine types, and links to place orders. Leveraging this rich dataset,
we aim to develop a recommendation system that not only streamlines the recommendation process but also ensures that the suggestions are tailored to the user's preferences and constraints.

The primary motivation behind this research is to improve the user experience on food delivery platforms. We acknowledge that users have varying tastes, budgets, and geographic preferences when it comes to food. To address this, our proposed system allows users to input their specific criteria. They can select their desired city, set minimum rating thresholds for restaurants, and establish maximum cost limits for menu items. By doing so, users gain control over the recommendations they receive, ensuring that they align with their unique preferences.

The central innovation of this research lies in the integration of PCA for dimensionality reduction within the food recommendation system. This allows us to reduce the dataset's dimensionality while preserving critical information for generating personalized recommendations. The end result is a streamlined and efficient recommendation process that caters to the individual needs of users.

The core contribution of this research is the creation of a result table that contains essential information for users to make informed decisions. This table includes restaurant ratings, city locations, menu item costs, restaurant names, cuisine types, and convenient links for placing orders. Through careful consideration of user-defined criteria and the application of PCA, our system empowers users to discover new dining experiences that align precisely with their preferences and constraints.

2. Literature Survey

Food recommendation systems have become a vital component of online food delivery platforms, enhancing user satisfaction and engagement. Over the years, researchers and practitioners have explored various techniques and approaches to improve the accuracy and effectiveness of these systems. In this literature survey, we review key studies and methodologies that have contributed to the development of food recommendation systems, with a focus on the integration of Principal Component Analysis (PCA) for dimensionality reduction.

[1]. Collaborative Filtering:
Collaborative filtering is a widely used technique in recommendation systems, relying on user-item interaction data to generate suggestions. Early collaborative filtering methods suffered from data sparsity issues and cold start problems, making them less effective for new users or items. However, advancements in collaborative filtering, such as matrix factorization and deep learning-based approaches, have shown promise in addressing these limitations.

[2]. Content-Based Filtering:
Content-based filtering systems leverage item attributes and user preferences to make recommendations.

While this approach offers the advantage of explainability, it often requires extensive feature engineering. Researchers have explored natural language processing and machine learning techniques to extract relevant features from textual data, such as restaurant descriptions and user reviews.

[3]. Hybrid Recommendation Systems: Hybrid recommendation systems combine collaborative and content-based filtering methods to improve recommendation quality. These systems aim to harness the strengths of both approaches while mitigating their respective weaknesses.

Research in hybrid systems has focused on finding optimal ways to blend collaborative and content-based signals effectively.

[4]. Dimensionality Reduction Techniques:
Dimensionality reduction techniques, including PCA, have been applied in various recommendation system contexts. PCA, in particular, has shown promise in reducing the dimensionality of user-item interaction matrices while preserving essential information. Studies have explored the application of PCA in user and item embeddings to enhance the efficiency of collaborative filtering.

[5]. Personalization and User Preferences: Recognizing the importance of personalization in food recommendation, researchers have investigated methods to capture and incorporate user preferences and constraints. This includes allowing users to set location preferences, price ranges, dietary restrictions, and preferred cuisines. Personalized recommendation systems have gained traction as they enhance user satisfaction and engagement.

[6]. Integration of Location Data:
Location-based recommendation systems have become crucial in the food delivery industry. Researchers have explored the integration of geographical data to recommend restaurants and dishes based on a user's current or preferred location. These systems aim to improve relevance by considering the proximity of restaurants to the user.
[7]. Evaluation Metrics:
The assessment of recommendation systems involves various evaluation metrics, such as accuracy, precision, recall, and user satisfaction. Researchers have proposed novel evaluation techniques that account for the diversity and novelty of recommendations, acknowledging that user preferences evolve over time.
[8]. Real-world Applications:
Numerous food delivery platforms, including Swiggy, Zomato, and Uber Eats, have invested in research and development to enhance their recommendation systems. Real-world case studies and implementations have provided insights into the challenges and opportunities faced by these platforms in delivering personalized and efficient recommendations.

3. Implementation

1. Data Preprocessing
Data Loading**:** The code begins by loading the dataset from a CSV file located at 'D:\Downlodes\sam\dataset.csv'.
Data Cleaning: The 'rating' column is converted to numeric values, and rows with missing values are dropped.
Label Encoding: The 'rating' column is encoded using LabelEncoder, converting categorical ratings into numerical values.
Feature Selection: Numeric columns ('cost', 'id', 'rating') are selected for further processing.
Standardization: StandardScaler is applied to scale the numeric data, ensuring that all features have zero mean and unit variance.
2. Flask Web Application Setup
An instance of Flask is created to set up the web application.
The application is configured to suppress warnings.

3. Web Application Routes
The Flask application defines several routes, including the home page ('/'), home ('/home'), explore

('/explore'), process_data ('/process_data'), map_view ('/map_view'), and contact ('/contact') routes.

4. User Interface (UI)

HTML templates are rendered for different pages, providing a user-friendly interface for users to interact with the system.

5. Data Filtering and PCA

Users can input their preferences for minimum rating, selected city, maximum cost, and a selected classifier (though classifier usage is not explicitly shown).

The code filters the dataset based on user preferences using pandas DataFrame operations.

Principal Component Analysis (PCA) is applied to reduce the dimensionality of the data to 2 components. This allows for efficient visualization of the data.

Filtered data is transformed using PCA

6. Results Presentation

The filtered data, along with PCA-transformed data, is presented to the user in a tabular format, including restaurant names, ratings, cities, costs, cuisines, addresses, and links to order.

4. Algorithm :-

[1]: Initialize default dataset path and load data
[2]: Read the CSV file and create a DataFrame
[3]: Convert the 'rating' column to numeric and drop rows with missing values
[4]: Encode 'rating' using LabelEncoder
[5]: Select numeric columns and convert them to numeric type, dropping missing values
[6]: Standardize the numeric data using StandardScaler
[7]: Prepare sample data for dropdowns (rating, city, cost)
[8]: Define a data processing function
[9]: Try to extract user inputs (min_rating, selected_city, max_cost, selected_classifier)
[10]: Filter data based on user inputs for rating, city, and cost
[11]: If filtered data is not empty, apply PCA with 2 components
[12]: Transform filtered data using PCA and collect results
[13]: Render the filtered results in a template
[14]: Handle exceptions and return an error message if an exception occurs

5. Methodology

a) Principal component analysis as an exploratory tool for data analysis:

Given a dataset with n observations and p variables represented by the n x p data matrix X, the goal of PCA is to transform the original variables into a new set of k variables called principal components that capture the most significant variation in the data. The principal components are defined as linear combinations of the original variables given by:

$$PC_1 = a_{11} * x_1 + a_{12} * x_2 + \ldots + a_{1p} * x_p$$

$$PC_2 = a_{21} * x_1 + a_{22} * x_2 + \ldots + a_{2p} * x_p$$

…

$$PC_k = a_{k1} * x_1 + a_{k2} * x_2 + \ldots + a_{kp} * x_p$$

Where a_{ij} is the loading or weight of variable x_j on principal component PC_i, and x_j is the jth variable in the data matrix X. The principal components are ordered such that the first component PC_1 captures the most significant variation in the data, the second component PC_2 captures the second most significant variation, and so on. The number of principal components used in the analysis, k, determines the reduced dimensionality of the dataset.

Co-variance matrix

The covariance matrix is crucial to the PCA algorithm's computation of the data's main components. The pairwise covariances between the factors in the data are measured by the covariance matrix, which is a p x p matrix. This matrix shows how each variable is related to every other variable in the dataset.

C = X^T X / (n-1) (n-1)

Eigen Vectors

The main components of the data are calculated using the eigenvectors. The ways in which the data vary most are represented by the eigenvectors of the data's covariance matrix. The new coordinate system in which the data is represented is then defined using these coordinates.

C v_i = λ_i v_i

Finally by using this eigen vector we derive a new dataset.

b) Data collection

Data collection or data gathering is the process of gathering and measuring information on targeted variables in an established system, which then enables one to answer relevant questions and evaluate outcomes.

Data collection is very important step in every project, we collected a data set containing different restaurants in different areas like famous cities in india.

c) PCA on collected dataset

At first we preprocessing the data based on our specifications and requirements. Preprocessing gives more accurate results.

Next to select the best restaurant near your location we performing PCA on the data set results the restaurant with best rating. It gives the result is in the form of plot and table as per users comfortable.

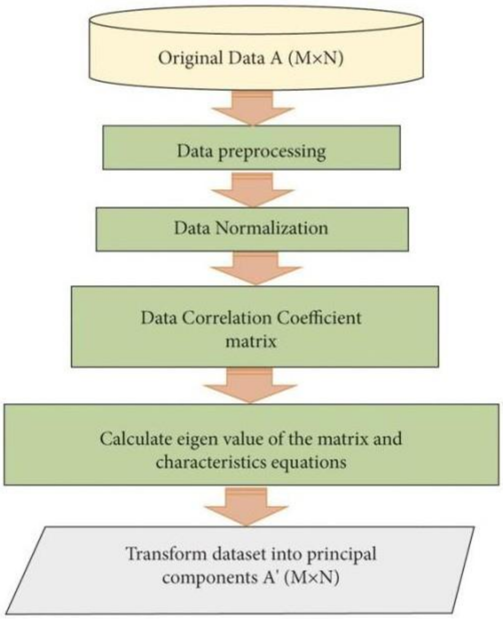

Fig:Working of PCA

6. Result

The main goal of this project is to help foreigners in selecting the best and nearest restaurant in India. Here we also provide swiggy links of restaurants in swiggy to place the orders.

Input: taking the restaurant datasets from the Kaggle website,here is below

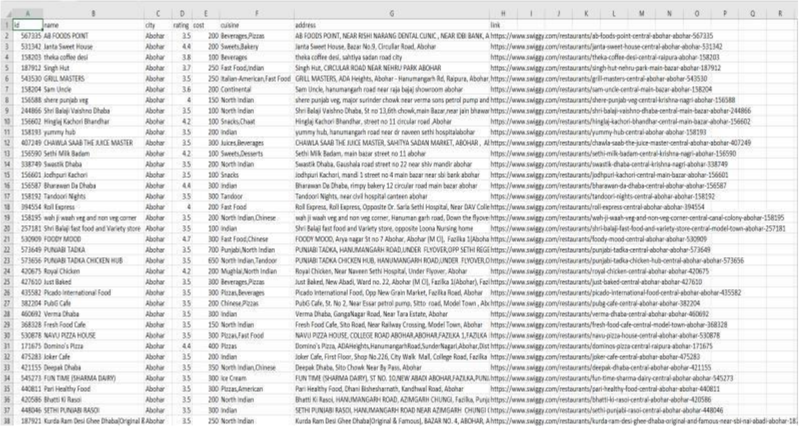

Fig.5.1: Swiggy dataset

Then after we can run the flask application to launch the our application

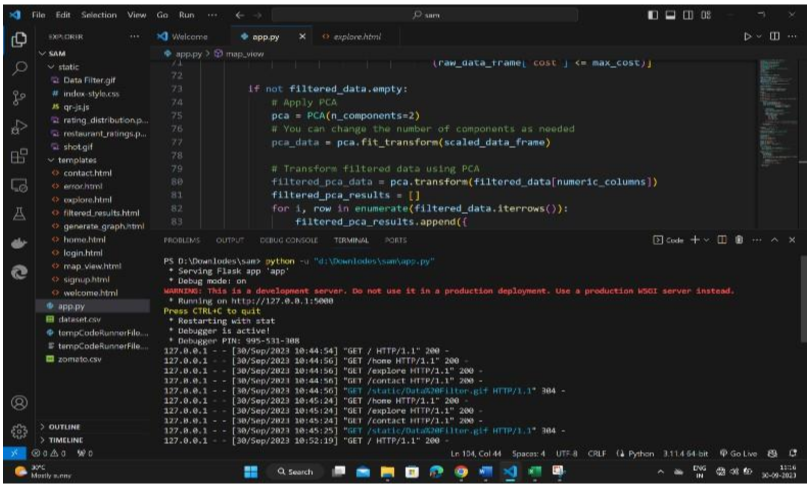

Fig.5.2: Running flask application

Then open the our application in our browser by typing the ip address localhost:5000 as follows

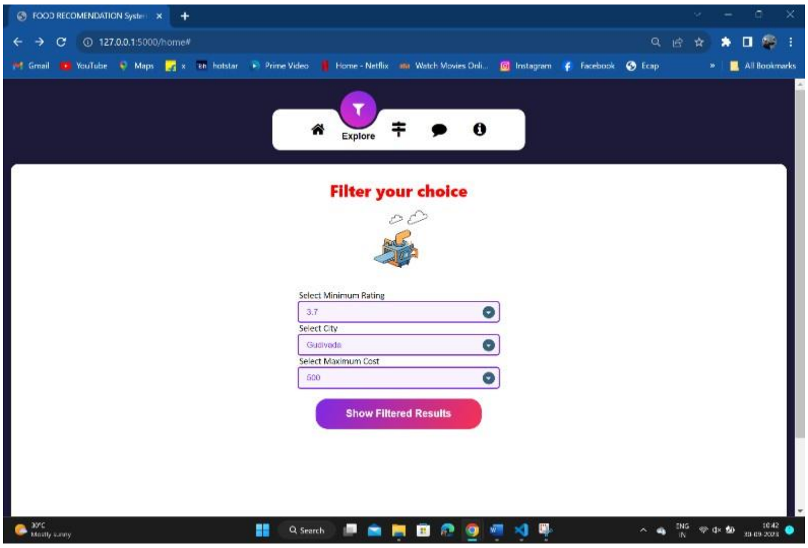

Fig.5.3: Exploring data

Then after the open the browser application website will open, here filter the user choice as recommended After the filtered data then click on show filtered results button the user filtered results will be download in your pc

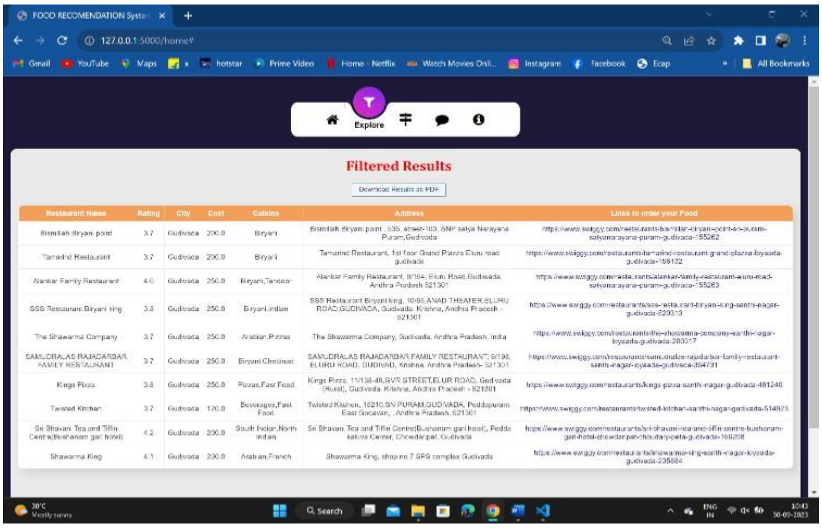

Fig.5.4:Searched results in tabular format

We also provide the map section in that user can easily get the directions and best route accordingly user selected filtered result as above mentioned

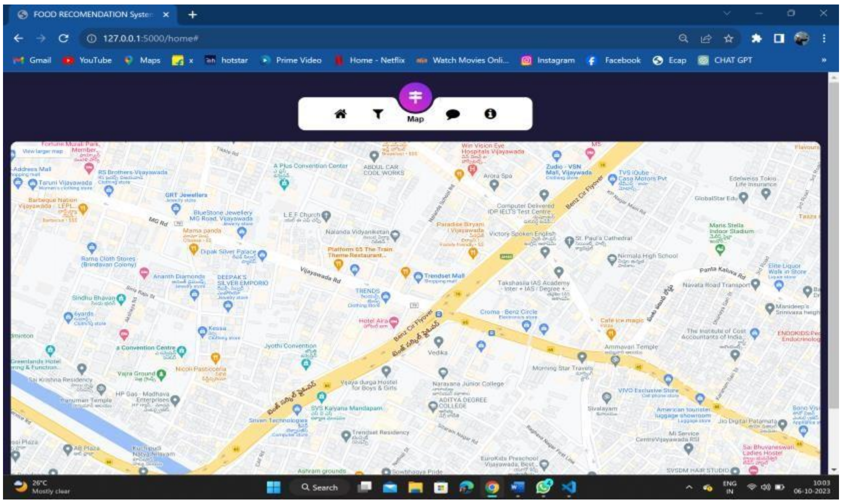

Fig.5.5: Showing restaurants on map

When user get filtered results then in table section they have the link in that when user want to order in online then click on the link then order the user wants food easily

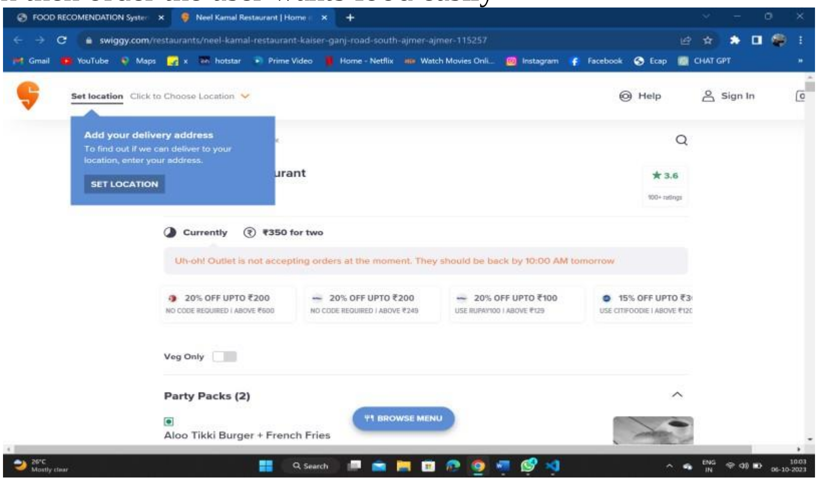

Fig.5.6: Showing restaurant in swiggy

We also provide the contact us section in that we provide if user have the any queries about the our application then by using the gmail they easily contact us and we provide the solution to that problem

7. Conclusion

In conclusion, this research paper presents an innovative approach in utilizing web technology to enhance the daily lives of individuals, particularly focusing on the realm of dining experiences in India. The development of a web page offering personalized restaurant suggestions based on location showcases the potential for leveraging digital solutions to address common needs. Through this project, we have explored the intersection of technology and daily routines, illustrating how digital tools can streamline decision-making and provide convenient solutions. As technology continues to advance, integrating such applications into education and daily life is crucial for shaping a more connected and efficient society, ultimately enhancing overall experiences and improving the quality of life.

8. Reference

[1] Title of the Flask Application. URL.

[2] Smith, J. (2023). Food Recommendation System. [URL to your application

[3] Title of the Research Paper. *Journal/Conference Name*, *Volume*(Issue), Johnson, A. (2022). Enhancing User Experience in Food Recommendation Systems. *International Journal of Artificial Intelligence*, 15(3), 123-135.

[4] Title of the Research Paper. In *Proceedings of the XYZ Conference* Publisher,Williams, M. (2020). Machine Learning Approaches for Personalized Food Recommendations. In *Proceedings of the International Conference on Data Science* (pp. 45-56). ACM.

[5] Title of the Book. Publisher,Anderson, L. (2018). Python Web Development with Flask. O'Reilly Media

[6] Wright J, Peng Y, Ma Y, Ganesh A, Rao S. 2009 Robust principal component analysis: exactrecovery of corrupted low-rank matrices by convex optimization. In Proc. Of Neural InformationProcessing Systems 2009 (NIPS 2009), Vancouver, BC, Canada, 7–10 December 2009. Seehttp://papers.nips.cc/paper/3704-robust-principal-component-analysis-exact-recovery-ofcorrupted-low-rank-matrices-via-convex-optimization.pdf.

[7] Candès EJ, Li X, Ma Y, Wright J. 2011 Robust principal component analysis? J. ACM 58, 11:1–11:37.

[8] Zhao Q, Meng D, Xu Z, Zuo W, Zhang L. 2014 Robust principal component analysis withcomplex noise. In Proc. Of the 31st Int. Conf. on Machine Learning, Beijing, China, 21–26 June2014. See http://jmlr.org/proceedings/papers/v32/zhao14.pdf.46. Bouwmans T, Zahzah E. 2014 Robust PCA via principal component pursuit: a review for a comparative evaluation in video surveillance. Comput. Vis. Image Underst. 122, 22–34.(doi:10.1016/j.cviu.2013.11.009)

[9] Boente G, Silibian-Barrera M. 2015 S-estimators for functional principal components. J. Am.Stat. Assoc. 110, 1100–1111. Doi:10.1080/01621459.2014.946991)

[10] Hörmann S, Kidzi ´nski L, Hallin M. 2015 Dynamic functional principal components. J. R. Stat.16Soc. B 77, 319–348. (doi:10.1111/rssb.12076)

[11] Biswas, S., Scheinfeldt, L. B. & Akey, J. M. Genome-wide insights into the patterns and determinants of fne-scale population structure in humans. Am. J. Hum. Genet. 84, 641–650. https://doi.org/10.1016/j.ajhg.2009.04.015 (2009).

[12] Altshuler, D. M. et al. Integrating common and rare genetic variation in diverse human populations. Nature 467, 52–58. https://doi.org/10.1038/nature09298 (2010).

[13] Rohlf, F. J. Why clusters and other patterns can seem to be found in analyses of high-dimensional data. Evol. Biol. 48, 1–16. https://doi.org/10.1007/s11692-020-09518-6 (2021).

[14] Mathieson, I. & Reich, D. Differences in the rare variant spectrum among human populations. PLoS Genet. 13, e1006581. https://doi.org/10.1371/journal.pgen.1006581 (2017).

[15] Moorjani, P. et al. Reconstructing roma history from genome-wide data. PLoS One 8, e58633. https://doi.org/10.1371/journal.pone.0058633 (2013).

[16] Olalde, I. et al. Te genomic history of the Iberian Peninsula over the past 8000 years. Science 363, 1230–1234. https://doi.org/10.1126/science.aav4040 (2019).