

Analyse internationaler Fußballergebnisse von 1872 bis 2024

Von Patrick Eicher, Demian In den Birken und Rick Springer

Über den Datensatz

Inhalt




Dieser Datensatz enthält 45.640 Ergebnisse internationaler Fußballspiele. Beginnend mit dem allerersten offiziellen Spiel im Jahr 1872 bis zum Jahr 2024. Die Spiele reichen vom FIFA World Cup über den FIFA Wild Cup bis hin zu regulären Freundschaftsspielen. Bei den Spielen handelt es sich ausschließlich um volle Länderspiele der Männer, Olympische Spiele oder Spiele, bei denen mindestens eine der Mannschaften teilgenommen hat. Die B-Mannschaft, die U-23 oder eine Auswahlmannschaft der Liga war bzw. sind in den Daten nicht enthalten.

Eine Übersicht der Daten und ergänzende Hinweise wird auf Seite 2 gegeben. (*Originaldaten von Mart Jürisod (Zuletzt aktualisiert 26.02.2024)*)

Fragestellungen

Innerhalb des Praxisprojekts werden folgende Leitfragen behandelt:

(● stehen für Mindestanforderungen, ○ stehen für wünschenswerte Ergänzungen)

- Wer ist die beste Mannschaft aller Zeiten? 
- Welche Mannschaften dominierten die verschiedenen Epochen des Fußballs?
- Welche Auffälligkeiten lassen sich im Laufe der Zeit beobachten?
 - Lassen sich Phänomene wie 'Heimvorteil', 'Auswärtsstärke', insgesamt erzielte Tore, Verteilung der Mannschaftsstärke usw. beobachten?
- Gab es Zeiten, in denen wenig oder gar keine Fußballspiele stattfanden? 
 - Wann fanden die meisten, wann die wenigsten Spiel statt?
 - Kann man anhand von Fußballspielen etwas über die Geopolitik aussagen - wie hat sich die Anzahl der Länder verändert, welche Mannschaften spielen oft gegeneinander?
- Welche Länder richten die meisten Spiele aus?
 - Welche Länder richten die meisten Spiele aus, an denen sie selbst nicht beteiligt sind?
- Inwieweit, wenn überhaupt, trägt die Ausrichtung eines großen Turniers zu den Chancen eines Landes im Turnier bei?
- Welche Mannschaften sind am aktivsten bei Freundschaftsspielen und -turnieren 
 - Welchen Mannschaften helfen oder schaden Freundschaftsspiele?

results.csv enthält die folgenden Spalten:

date - Datum des Spiels

home_team - der Name der Heimmannschaft

away_team - Name der Auswärtsmannschaft

home_score - Gesamtergebnis der Heimmannschaft, einschließlich Verlängerung, ohne Elfmeterschießen

away_score - Gesamtergebnis der Auswärtsmannschaft, einschließlich Verlängerung, ohne Elfmeterschießen

tournament - der Name des Turniers

city - Name der Stadt/des Ortes/der Verwaltungseinheit, in der/dem das Spiel ausgetragen wurde

country - Name des Landes, in dem das Spiel ausgetragen wurde

neutral - TRUE/FALSE-Spalte, die angibt, ob das Spiel an einem neutralen Ort ausgetragen wurde

shootouts.csv enthält die folgenden Spalten:

date - Datum des Spiels

home_team - der Name der Heimmannschaft

away_team - Name der Auswärtsmannschaft

winner - Gewinner des Elfmeterschießens

first_shooter - die Mannschaft, die als erste ins Elfmeterschießen ging

goalscorers.csv enthält die folgenden Spalten:

goalscorers.csv enthält die folgenden Spalten:

date - Datum des Spiels

home_team - der Name der Heimmannschaft

away_team - Name der Auswärtsmannschaft

team - Name der Mannschaft, die das Tor geschossen hat

scorer - Name des Spielers, der das Tor geschossen hat

own_goal - ob das Tor ein Eigentor war

penalty - ob das Tor ein Elfmeter war

Hinweis zu Mannschaften- und Ländernamen:

Für Heim- und Auswärtsmannschaften wurde der aktuelle Name (15.03.2024) der Mannschaft verwendet. Wenn zum Beispiel 1882 eine Mannschaft, die sich Irland nannte, gegen England spielte, wird sie in diesem Datensatz als Nordirland bezeichnet, weil die aktuelle Mannschaft Nordirlands der Nachfolger der irischen Mannschaft von 1882 ist. Auf diese Weise ist es einfacher, die Geschichte und die Statistiken der Mannschaften zu verfolgen.

Für Ländernamen wird der Name des Landes zum Zeitpunkt des Spiels verwendet. Wenn also Ghana in den 1950er Jahren in Accra an der Goldküste spielte, war es ein Heimspiel für Ghana, auch wenn die Namen der Heimmannschaft und des Landes nicht übereinstimmen. Dies geht aus der Spalte "Neutral" hervor, in der für diese Spiele FALSCH steht, was bedeutet, dass das Spiel nicht an einem neutralen Ort stattfand.