



موضوع پروژه:

بررسی و تحلیل داده های مسابقات فوتبال با هدف پیدا کردن دلایل گل زدن و برد تیم ها

۱- هدف پروژه چیست؟

هدف از این پروژه این است که با توجه به داده های مسابقات فوتبال در تورنومنت های مختلف از جمله جام جهانی، یورو و ... که شامل داده های event based از جمله پاس، شوت، گل و ... است چه تفاوتی بین این event ها در زمان قبل از گل زدن و به طور کلی در کل زمان بازی برای تیم های برنده وجود دارد.

۲- منبع داده از کجاست؟

داده های به صورت آزاد در repository گیت هاب شرکت [statsbomb](https://statsbomb.com) در این [لینک](#) قرار دارد.

۳- چه سوالاتی باید پاسخ داده شوند؟

۱- تفاوت های زیر در زمان بازی فعال قبل از به ثمر رسیدن گل و بقیه زمان ها:

- تعداد پاس
- زمان مالکیت توپ
- تعداد بازیکنان نزدیک به دروازه
- فاصله شوت به دروازه
- ...

۲- کدام ویژگی ها از جمله میزان دوندگی، تعداد پاس، تعداد شوت و ... برای تیم های برنده متفاوت است.

۳- تفاوتی در توزیع پنالتهی زدن (چپ، راست، وسط) برای تیم های برنده و غیر برنده وجود دارد یا خیر.

۴- بررسی موضوعات دیگر بر اساس داده که مشخص شوند تیم های برنده در این ویژگی ها با تیم های بازنده متفاوت هستند.

۴- داده به چه شکل جمع آوری شده است؟

توضیحی در مورد نحوه جمع آوری داده ها در لینک قرار داده نشده است زیرا شرکت statsbomb محصولات غیر رایگان مختلفی در زمینه تحلیل داده های فوتبال دارد و این مجموعه داده ها را با هدف تحقیق و گسترش تحلیل داده های فوتبال قرار داده است.

۵- چه جدول هایی در داده وجود دارد؟

- جدول مسابقات (competitions): که اسم و اطلاعات مربوط به مسابقات (تورنومنت) های مختلف در آن وجود دارد.
- جدول بازی ها (matches): اطلاعات مربوط به هر بازی فوتبال در آن وجود دارد.
- جدول اتفاقات بازی (events): که در آن اتفاقات مختلف بازی از جمله گل، پاس، شوت و ... و اطلاعات مکانی و زمانی آن

۶- جزئیات جدول ها به چه شکلی است؟

ستون های دیگری در این جدول ها وجود دارد ولی ستون های اصلی و احتمالا مورد نیاز در زیر آورده شده اند.

competition:

competition_id, season_id, country_name, competition_name,
competition_gender, competition_youth, competition_international,
season_name

matches:

match_id, match_date, kick_off, competition, season, home_team, away_team, home_score, away_score, match_status, match_week

Events:

event_type, id, index, period, timestamp, minute, second, possession, possession_team, play_pattern, team, player, position, duration, start_location_x, start_location_y, end_location_x, end_location_y, end_location_z

۷- در صورتی که منبع داده پابلیک است لطفا لینک صفحه داده را بنویسید.

<https://github.com/statsbomb/open-data>