

# PROGETTO DI DATA ANALYST



PIERLUIGI VENNARI

## INTRODUZIONE

Il tennis è storicamente associato all'élite aristocratica come uno degli sport prediletti da re e nobili, il tennis è oggi un'appassionante disciplina che attrae individui di ogni estrazione sociale, età e background. La sua popolarità è alimentata dalla semplicità delle regole, dalla minima attrezzatura richiesta e dalla flessibilità che consente di praticarlo sia individualmente che in coppia. Attualmente, i tornei più prestigiosi sono gestiti da organizzazioni internazionali di spicco, ovvero l'ATP (Association of Tennis Professionals) per il circuito maschile e la WTA (Women's Tennis Association) per quello femminile. Tra le competizioni di maggior rilievo rientrano i celebri tornei del Grande Slam, tra cui l'Open d'Australia, Roland Garros, Wimbledon e lo U.S. Open, seguiti dai tornei Masters 1000, 500 e 250.

## PROGETTO DAPT0123 - ACEDATASPORT

Nel vasto universo dell'analisi dati, poche discipline offrono un palcoscenico così affascinante e dinamico quanto il mondo del tennis. Il tennis è uno sport intriso di tradizione, ma allo stesso tempo in continua evoluzione. È un connubio di talento, strategia, e una dose di imprevedibilità che lo rendono affascinante per gli appassionati di sport di tutto il mondo. Ma cosa accade quando uniamo l'incredibile potenza dell'analisi dati a questo sport affascinante? Nasce una sinergia straordinaria, una fusione di dati e competenze che ci permette di svelare nuove dimensioni dell'esperienza tennistica. Benvenuti in questo progetto di analisi dati nel tennis, un'esplorazione avvincente delle statistiche e delle dinamiche di questo sport, supportata da potenti strumenti come Microsoft Power BI.



Figura 1 - Home progetto

In questo viaggio, ci immergeremo nelle cifre e nei fatti che circondano il tennis professionistico, scrutando da vicino la differenza di prestazioni tra uomini e donne, e mettendo in luce l'essenza stessa delle partite e dei giocatori attraverso grafici informativi e tabelle statistiche appositamente progettate. L'uso di Power BI, un software di visualizzazione dati all'avanguardia, ci permetterà di tradurre i dati complessi in visualizzazioni chiare e significative. Potremo esplorare le tendenze, identificare modelli e rivelare dettagli che sfuggono all'occhio non allenato. Attraverso questa potente piattaforma, avremo la possibilità di analizzare in profondità i dati del tennis, fornendo una panoramica completa delle prestazioni di giocatori e giocatrici, così come delle dinamiche delle partite. Una delle sfide affrontate in questo progetto sarà l'analisi comparativa tra uomini e donne nel tennis professionistico. Oltre alle evidenti differenze fisiche, ciò che emerge dai dati può rivelare aspetti intriganti sulle strategie di gioco, la consistenza delle performance e la crescita del tennis femminile nell'ambito del professionismo. Le tabelle statistiche, create appositamente per questo progetto, rappresenteranno uno strumento fondamentale per estrarre informazioni dettagliate e sintetizzare i dati in modo chiaro e accessibile. Saranno un compagno essenziale per comprendere l'andamento delle partite, le statistiche dei giocatori e le tendenze nel tennis moderno.



*Figura 1- Logo progetto*

## **1. ANALISI DEI MIGLIORI GIOCATORI ATP E WTA**

Questo Report consente di confrontare i migliori giocatori ATP e WTA in diverse statistiche. Si trovano, statistiche sui Game, partite, Ace, set vinti e breakpoint guadagnati. Vengono esaminati due fonti che contengono i 50 migliori tennisti:

- Uomini: <http://www.tennisabstract.com/cgi-bin/leaders.cgi>
- Donne: [http://www.tennisabstract.com/cgi-bin/leaders\\_wta.cgi](http://www.tennisabstract.com/cgi-bin/leaders_wta.cgi)

Ho trasformato questi dati in file CSV e li ho importati in Power BI. L'obiettivo principale era esaminare la possibile esistenza di diversi "profili" di giocatori di tennis e valutare come tali profili potessero essere correlati al genere dei partecipanti. E' importante notare che il set di dati è composto da 100 giocatori, e comprende principalmente dati relativi alle prestazioni (ad esempio, punti vinti al servizio), mentre mancano informazioni dettagliate sugli aspetti fisici del gioco (come la velocità del servizio). È evidente che gli uomini superano le donne in termini di punti guadagnati al servizio, ace realizzati e giochi vinti al servizio. Questa disparità è comprensibile, considerando che gli uomini tendono probabilmente a possedere servizi più potenti e difficili da contrastare. È altresì plausibile che le donne abbiano focalizzato maggiormente il loro sviluppo nel miglioramento delle abilità di risposta, mentre gli uomini abbiano dato maggiore attenzione all'incremento delle prestazioni nel servizio.

## 2. ANALISI DEGLI ULTIMI 20 ANNI DEL GRANDE SLAM

Esistono quattro tornei di primissimo piano noti come il "Grande Slam," ognuno disputato su tre diverse superfici: terra battuta, cemento e erba. Lo scopo di questa analisi è l'approfondimento dei dati relativi agli ultimi 20 anni dei tornei del Grande Slam, concentrandosi specificamente su Wimbledon, l'Open d'Australia, lo US Open e il Roland Garros, ma limitando l'analisi all'ATP. Il set di dati utilizzato è composto da tre diverse raccolte di informazioni:

1-Match.csv: Questo file contiene una vasta gamma di dati relativi alle singole partite, tra cui informazioni come il turno del torneo, il nome del torneo, l'anno in cui si è svolto, e altro ancora.

2-Player.csv: Questo dataset comprende tutti i dettagli rilevanti di ogni singolo giocatore partecipante, tra cui data di nascita, mano dominante di gioco, nazionalità e altri attributi.

3-Stats.csv: In questo file sono raccolti tutti i dati relativi alle performance di ciascun giocatore nelle partite disputate, tra cui il numero di ace, i punti guadagnati, i break point vinti o salvati, la classifica e molti altri indicatori significativi.

Questo progetto mira a esaminare in modo approfondito le sfumature delle prestazioni dei giocatori di tennis in questi prestigiosi tornei del Grande Slam, fornendo una panoramica completa delle sfide e dei trionfi che hanno caratterizzato gli ultimi 20 in queste importanti competizioni.