Iron Hack | Data Analytics

Poker Playbook Analysis:
A Comparative Study of Winner and Losing players

Presentation

João Miguel Cortez Sobral

Poker Hand Rankings

Royal Flush:
Ace, King, Queen, Jack, and Ten of the same suit.

Straight Flush:
Five consecutive cards of the same suit.

Four of a Kind:
Four cards of the same rank.

Full House:
Three cards of one rank and two cards of another rank.

Flush:
Five cards of the same suit, not in sequence.

Straight:
Five consecutive cards of different suits.

Three of a Kind:
Three cards of the same rank.

Two Pairs:
Two sets of two cards with the same rank.

One Pair:
Two cards of the same rank.

High Card:
The highest card in the hand

Poker terminology

Limp
Calling the Minimum
Bet Without Raising

Pre flop raise(pfr)
Increasing the Bet

Big blind(bb)
big
mandatory
fee

3 bet
Re-Raising
After a raise

Small blind(sb)
small
mandatory
fee

4 bet
Re-Raising
After 2 raises

bb/100:
Measures winnings per 100 hands
1bb/100 = 0.20 cents per 100 hands
10bb/100 = 2 euros per 100 hands
- 10bb/100 = - 2 euros per 100 hands

Voluntarily put money in pot(VPIP)
Percentage of Hands a Player plays

Background and Methodology

01

Project Details

The primary goal of this project was to dive into players behavior in online poker, focusing on the NL 25 room of PokerKing.
By analyzing strategies employed during gameplay, we aimed to uncover insights into winning and losing player dynamics.

02

Data Details

We acquired a substantial dataset of 1.5 million poker hands from HH dealer.
This raw data was then transformed into a usable format through PokerTracker, facilitating its analysis. The dataset was further refined and processed using Jupyter Notebook and SQL to draw meaningful conclusions.

03

Hypothesis testing

Our project's hypothesis centered around understanding which specific gameplay statistics had the most influence on winnings and losses.
Through rigorous statistical testing, we aimed to pinpoint these influential factors and establish their significance in player performance.

04

Logistic regression

To gain a deeper understanding of player types, we classified participants as either "sharks" (successful players) or "fishes" (less successful players) based on their in-game winnings.
Utilizing logistic regression, we modeled the relationships between various statistics and player classification, imrpoving our insights into the distinct strategies adopted by different player categories.

Key findings on hypothesis testing

Null Hypothesis (H0): The means of all the statistics are the same for "Fish" players compared to "Shark" players.

Alternative Hypothesis (Ha): At least one of the means of all the statistics are different for "Fish" players compared to "Shark" players.

With a 95% confidence level, **the results indicated that the means of several statistics significantly differ between "Fish" and "Shark"** players therefore **we reject H0 and confirm the Ha**:(At least one of the means of all the statistics are different for "Fish" players compared to "Shark" players.)

Sharks(winning players)

-Play less hands in general (VPIP mean : 24.8%)

-Call less hands (Call any preflop raise mean: 13.5%)

-Make less limps (Limp mean: 1.1%)

-Are more agressive when playing
(3 bet mean: 8.7%)(4bet mean: 7.65%)

-Give up more versus agression when playing
(Fold to 3 bet mean: 75%) (Fold to 4 bet + mean: 60%)

Fishes(Losing players)

-Play more hands in general (VPIP mean : 28.8%)

-Call more hands (Call any preflop raise mean:21.4%)

-Make more limps (Limp mean: 4.5%)

-Are less agressive when playing
(3 bet mean:7.7%)(4bet mean: 6.2%)

-Give up less versus agression when playing
(Fold to 3 bet mean: 70.5%) (Fold to 4 bet + mean: 56.8%)

Diving deeper into findings

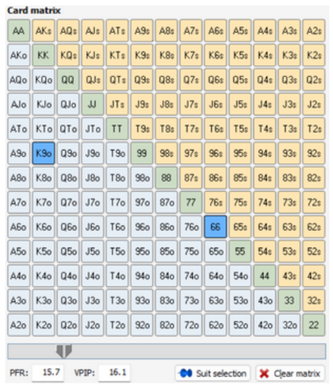
Sharks(winning players)

-Play less hands in general (VPIP mean : 24.8%)
-Call less hands (Call any preflop raise mean: 13.5%)
-Make less limps (Limp mean: 1.1%)
-Are more agressive when playing
(3 bet mean: 8.7%)(4bet mean: 7.65%)
-Give up more versus agression when playing
(Fold to 3 bet mean: 75%) (Fold to 4 bet + mean: 60%)

Avg Played hands by sharks (VPIP)

Avg call any preflop raise by sharks(cpfr))

Avg limp by sharks



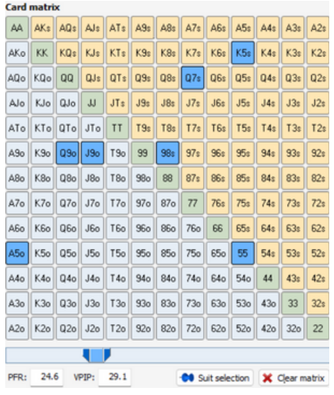
Fishes(Losing players)

-Play more hands in general (VPIP mean : 28.8%)
-Call more hands (Call any preflop raise mean:21.4%)
-Make more limps (Limp mean: 4.5%)
-Are less agressive when playing
(3 bet mean:7.7%)(4bet mean: 6.2%)
-Give up less versus agression when playing
(Fold to 3 bet mean: 70.5%) (Fold to 4 bet + mean: 56.8%)

Avg Played hands by fishes (VPIP)

Avg call any preflop raise by fishes(cpfr))

Avg limp by fishes



List of hands that Fishes play and sharks don't:

List of hands that Fishes call and sharks don't:

List of hands that Fishes limp and sharks don't:

Logistic regression

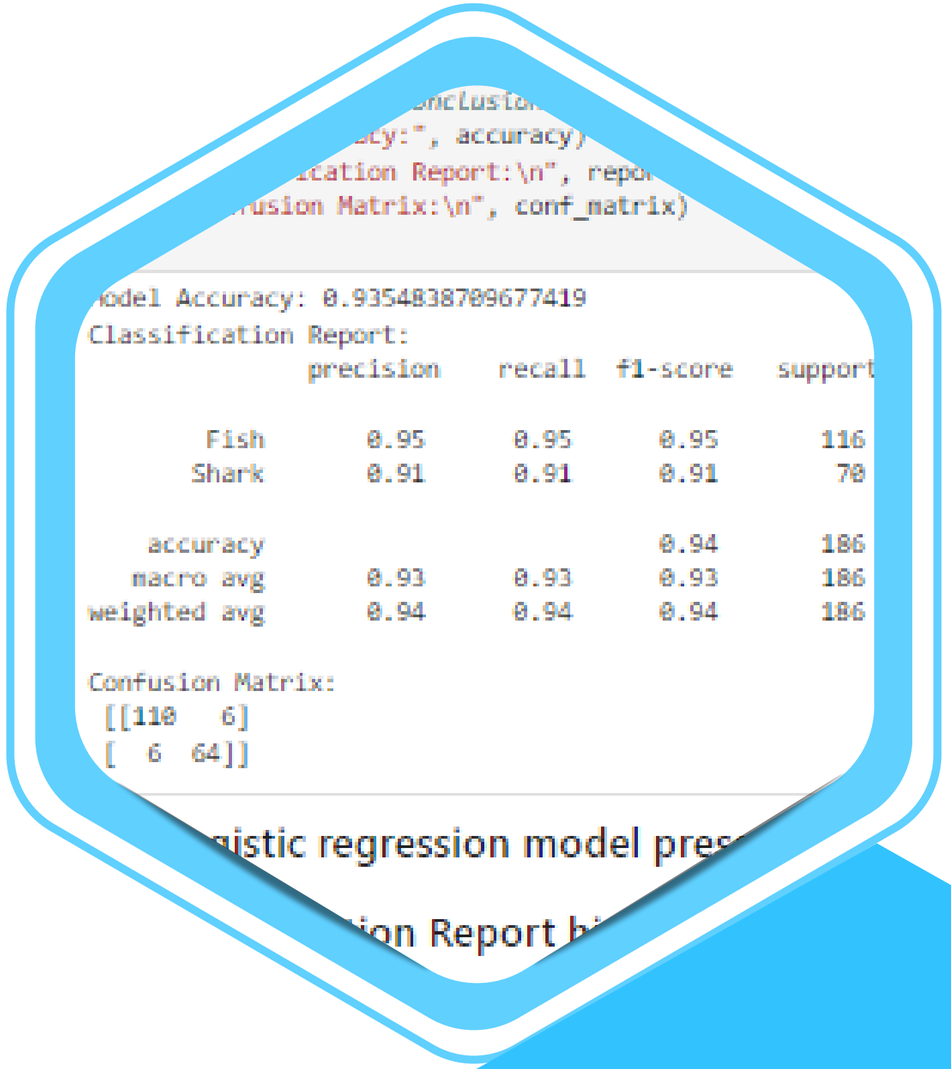
Our logistic regression model was designed to predict player types (Fish or Shark) using significant gameplay statistics. The model achieved a high accuracy of 93.55%, showcasing its ability to classify players.

Precision, recall, and F1-score metrics further validate its performance. Fish players were classified with a precision, recall, and F1-score of 0.95 each, while for Shark players, these metrics were at 0.91.

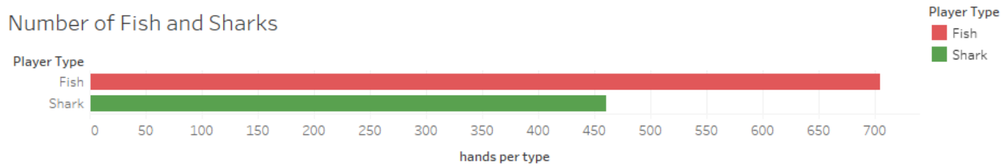
By selecting statistics with proven influence on winnings through hypothesis testing, we crafted a model capable of differentiating good players from bad players.
These insights allowed us to create a robust classification model.

The detailed confusion matrix revealed that only 6 instances each of sharks and fishes were misclassified, highlighting the model's reliability in player type prediction.

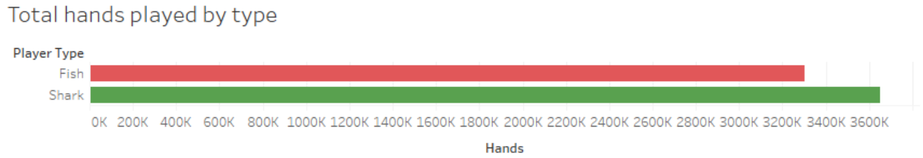
In summary, our logistic regression model is a potent tool for predicting player types based on gameplay statistics. This predictive capability allows us to understand player behavior and refine poker strategies with actionable insights.



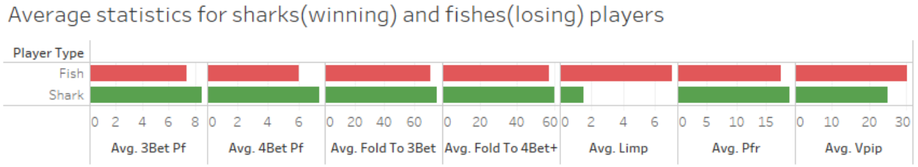
Data Visualizations with Tableau: Insights through Interactive Graphics

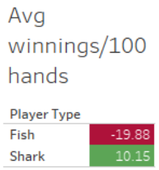


Number of player by type:
Fishes: 704
Sharks:460
Total players: 1164



Total hands played by type:
Fishes: 3 301 180
Sharks:3 652 707
Total hands played: 6 953 887





Average winnings bb/100 by type:
Fishes: -19.88 == -(0.25 x 19.88) == Lose 4.97$ per 100 hands on average
Sharks: 10.15 == 0.25 +10.15 == Win 2.54$ every 100 hands on average

Lab: Identifying fishes through statistics

Read Me: The goal of this lab is to identify if a player is likely to be a fish or a shark and state at least 3 reasons for your decision.

Next to each player , you will be able to see a HUD(Heads-up Display) a tool frequently used by poker players to visualize each player statistics.

Hints:

Fishes (Losing players) profiling according to our findings:

-Play more hands in general (bigger VPIP%)

-Call more hands (bigger Call any preflop raise%)

-Make more limps (bigger Limp%)

-Are less agressive when playing (smaller 3 bet%)(smaller 4bet%)

-Give up less versus agression when playing (smaller Fold to 3 bet %) (smaller Fold to 4 bet%)

Iron Hack | Data Analytics

The floor is open for questions!