Poker Hand History Analysis

Casey Tattersall

“Collecting Data”

- I collected a sample of 3,323 hands (through the power of multi-tabling).
  - Conventionally, you need 30,000 hands before you can draw any real conclusions, and more detailed analyses need well over 1,000,000
  - This sample is large enough to make our analysis interesting, but some of our results (specifically the most and least profitable hands) need to be taken with a grain of salt
- After playing, I downloaded an auto-generated hand history which contained a record of what happened in each hand, including the cards that all of my opponents were dealt.

```
Bovada Hand #4422641660 TBL#28825156 HOLDEM No Limit - 2023-02-16 19:34:10
Seat 1: Big Blind ($42.41 in chips)
Seat 2: UTG ($29.83 in chips)
Seat 3: UTG+1 ($18.38 in chips)
Seat 4: UTG+2 ($23.13 in chips)
Seat 5: Dealer [ME] ($37.39 in chips)
Seat 6: Small Blind ($24.75 in chips)
Dealer  [ME] : Set dealer [5]
Small Blind : Small Blind $0.10
Big Blind : Big blind $0.25
*** HOLE CARDS ***
Big Blind : Card dealt to a spot [Ah 6s]
UTG : Card dealt to a spot [7h 8h]
UTG+1 : Card dealt to a spot [9d Tc]
UTG+2 : Card dealt to a spot [9c Qd]
Dealer  [ME] : Card dealt to a spot [7c Kc]
Small Blind : Card dealt to a spot [6h 8d]
UTG : Calls $0.25
UTG+1 : Calls $0.25
UTG+2 : Calls $0.25
Dealer  [ME] : Raises $1.25 to $1.25
Small Blind : Folds
Small Blind : Table deposit $0.35
Big Blind : Calls $1
UTG : Calls $1
UTG+1 : Calls $1
UTG+2 : Folds
*** FLOP *** [5c 4h 8c]
Big Blind : Checks
UTG : Bets $2.55
UTG+1 : Folds
Dealer  [ME] : Calls $2.55
Big Blind : Calls $2.55
*** TURN *** [5c 4h 8c] [Qs]
Big Blind : Checks
UTG : Checks
UTG+1 : Seat stand
UTG+1 : Table leave user
Dealer  [ME] : Bets $8.25
Table enter user
Seat sit down
Table deposit $20
Big Blind : Folds
UTG : Calls $8.25
*** RIVER *** [5c 4h 8c Qs] [Ac]
UTG : All-in $17.78
Dealer  [ME] : Calls $17.78
UTG : Showdown [8h 8c Ac Qs 7h] (One pair)
Dealer  [ME] : Showdown [Ac Kc 8c 7c 5c] (Flush)
Dealer  [ME] : Hand result $63.06
UTG : Seat sit out
*** SUMMARY ***
Total Pot($65.06)
Board [5c 4h 8c Qs Ac]
Seat+1: Big Blind Folded on the TURN
Seat+2: UTG lost with One pair [7h 8h-8h 8c Ac Qs 7h]
Seat+3: UTG+1 Folded on the FLOP
Seat+4: UTG+2 Folded before the FLOP
Seat+5: Dealer $63.06  with Flush [7c Kc]
Seat+6: Small Blind Folded before the FLOP
```

Extracting the Data

- Each hand was recorded in a .txt file
  - The overall structure remained the same, but the order and number of lines changed significantly based on the action, number of players, positions, etc.
  - To automate the data collection process, we can leverage patterns in the formatting that are always present, and extract the relevant information with RegEx
- Example: Find the opener (first player, if any, to raise before the flop)

```
```{r}
str_match(str_match(hands[1652], "*** HOLE CARDS ***([^*]+)*")[2],
          "([^\r\n\\[\\]]{3,11}) (?: \\[ME\\] )?: (?:Raises|All\\-in)")[2]
```

[1] "Dealer"
```

- All of the data was stored in R6 classes, with a Hand class object for each hand and a Player class object for each player in the hand
  - We can then loop over all hands to extract the info we want to study into a DataFrame, and filter it with dplyr

Representing Poker Hands

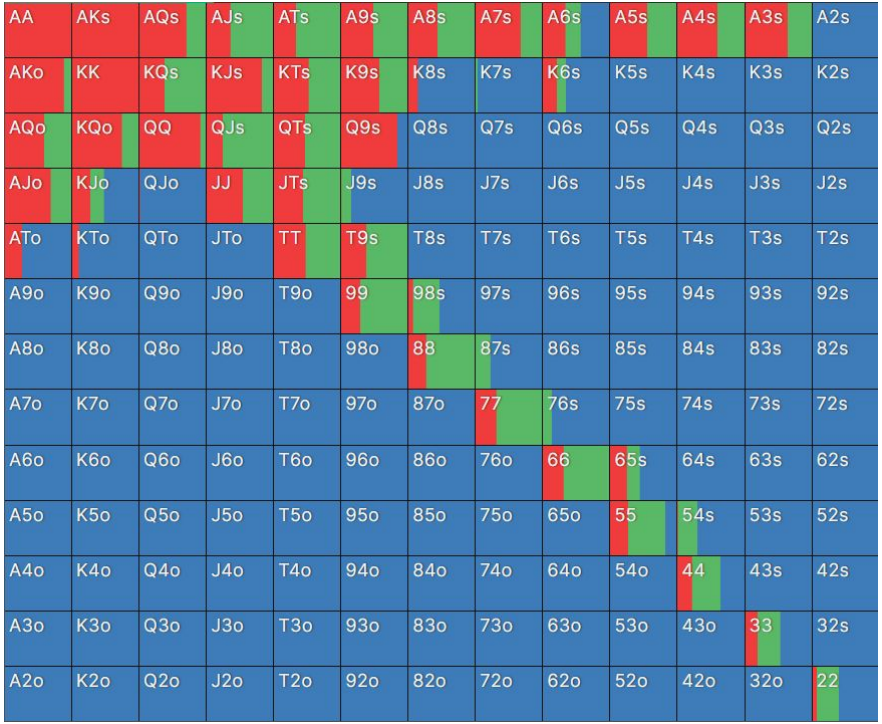
- In Texas Hold'em Poker, each player is dealt two cards, giving us 1,356 possible starting hands
- However, there is no ranking of the suits, so we can group hands together into suited (both cards are the same suit) and offsuit (the cards are different suits), along with pairs.
  - For example, an Ace and King of spades is effectively the same as an Ace and King of diamonds, so both are stored as AKs (where s = suited).
  - We still care about the suits because a flush (five cards of the same suit) is a really good hand, and if both of our starting cards are the same suit, we have a better chance of making one.
- This leaves us with 169 possible hands, which can be neatly arranged into a 13x13 grid of offsuit, paired, and suited hands.

AA	AKs	AQs	AJs	ATs	A9s	A8s	A7s	A6s	A5s	A4s	A3s	A2s
AKo	KK	KQs	KJs	KTs	K9s	K8s	K7s	K6s	K5s	K4s	K3s	K2s
AQo	KQo	QQ	QJs	QTs	Q9s	Q8s	Q7s	Q6s	Q5s	Q4s	Q3s	Q2s
AJo	KJo	QJo	JJ	JTs	J9s	J8s	J7s	J6s	J5s	J4s	J3s	J2s
ATo	KTo	QTo	JTo	TT	T9s	T8s	T7s	T6s	T5s	T4s	T3s	T2s
A9o	K9o	Q9o	J9o	T9o	99	98s	97s	96s	95s	94s	93s	92s
A8o	K8o	Q8o	J8o	T8o	98o	88	87s	86s	85s	84s	83s	82s
A7o	K7o	Q7o	J7o	T7o	97o	87o	77	76s	75s	74s	73s	72s
A6o	K6o	Q6o	J6o	T6o	96o	86o	76o	66	65s	64s	63s	62s
A5o	K5o	Q5o	J5o	T5o	95o	85o	75o	65o	55	54s	53s	52s
A4o	K4o	Q4o	J4o	T4o	94o	84o	74o	64o	54o	44	43s	42s
A3o	K3o	Q3o	J3o	T3o	93o	83o	73o	63o	53o	43o	33	32s
A2o	K2o	Q2o	J2o	T2o	92o	82o	72o	62o	52o	42o	32o	22

Representing Poker Hands

- We can use this 13x13 grid to represent how each hand should play (if discussing strategy) or did play (if reviewing a hand history) in a given scenario
  - Red = Raise (darker shade = bigger raise)
  - Green = Call
  - Blue = Fold

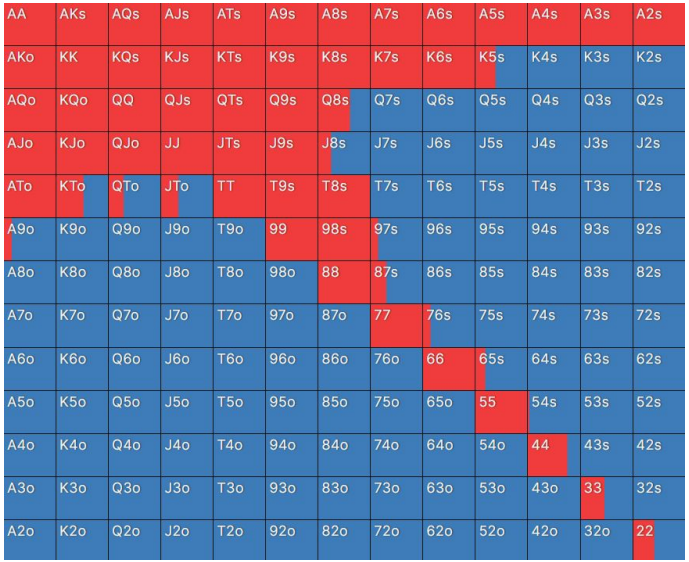
- If a hand is split between multiple actions, then proportional amounts of that box are shaded
  - For example, according to the solver, the hand **44** (a pair of fours) is raised about 20% of the time, called about 40% of the time, and folded about 30% of the time in this situation



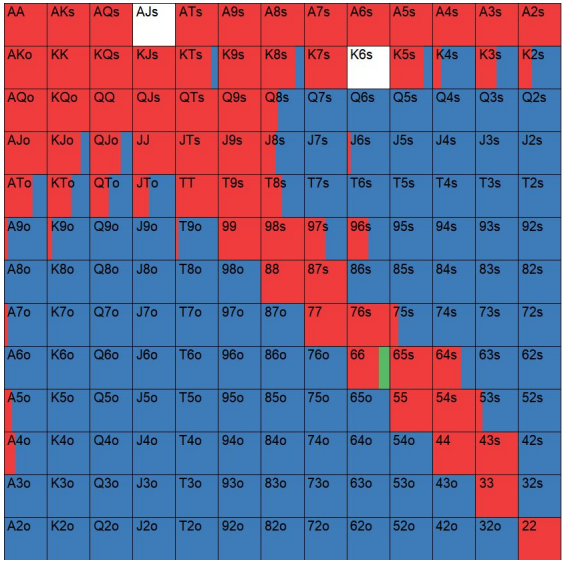
Preflop: BTN vs HJ open (GTOWizard Solver)

Comparison: Early Position Opening Range

Solver's solution ("correct" way of playing)



My actions (from the hand history)

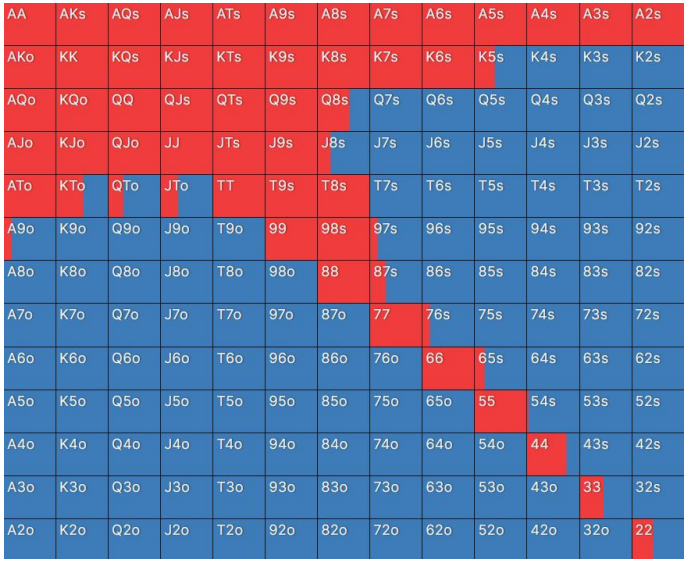


Pretty close! While I play a few too many low suited connectors (54s, 65s, etc), the charts are very similar.

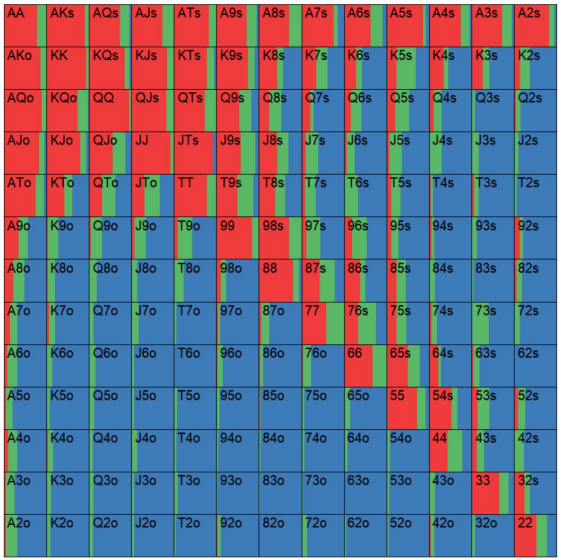
Note: Hands in white (AJs, K6s) were never dealt to me in this position in the sample.

Comparison: Early Position Opening Range

Solver's solution ("Correct" way of playing)



Opponent's actions (from the hand history)



These are way different - while the underlying raising pattern is the same, there's a lot of added noise. Specifically, just calling preflop (i.e. limping) is universally considered to be bad in this scenario, but inexperienced players still do it a lot.

Most and Least Profitable Hands - Me

AA	AKs	AQs	AJs	ATs	A9s	A8s	A7s	A6s	A5s	A4s	A3s	A2s
AKo	KK	KQs	KJs	KTs	K9s	K8s	K7s	K6s	K5s	K4s	K3s	K2s
AQo	KQo	QQ	QJs	QTs	Q9s	Q8s	Q7s	Q6s	Q5s	Q4s	Q3s	Q2s
AJo	KJo	QJo	JJ	JTs	J9s	J8s	J7s	J6s	J5s	J4s	J3s	J2s
ATo	KTo	QTo	JTo	TT	T9s	T8s	T7s	T6s	T5s	T4s	T3s	T2s
A9o	K9o	Q9o	J9o	T9o	99	98s	97s	96s	95s	94s	93s	92s
A8o	K8o	Q8o	J8o	T8o	98o	88	87s	86s	85s	84s	83s	82s
A7o	K7o	Q7o	J7o	T7o	97o	87o	77	76s	75s	74s	73s	72s
A6o	K6o	Q6o	J6o	T6o	96o	86o	76o	66	65s	64s	63s	62s
A5o	K5o	Q5o	J5o	T5o	95o	85o	75o	65o	55	54s	53s	52s
A4o	K4o	Q4o	J4o	T4o	94o	84o	74o	64o	54o	44	43s	42s
A3o	K3o	Q3o	J3o	T3o	93o	83o	73o	63o	53o	43o	33	32s
A2o	K2o	Q2o	J2o	T2o	92o	82o	72o	62o	52o	42o	32o	22

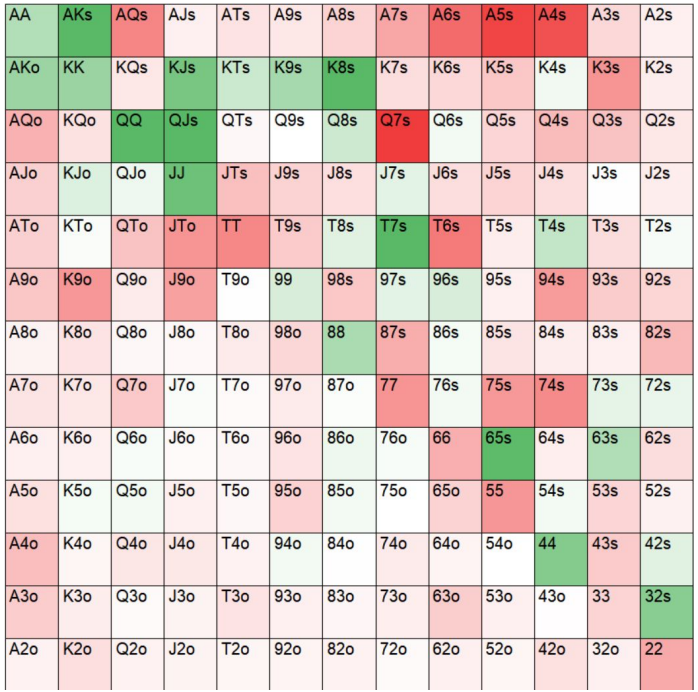
Green = net profit **Red = net loss**

- Five most profitable hands:
  - **JJ,** A9s, 87s, A5s, KQs
- Five least profitable hands:
  - **AKs**, 43s, 77, 55, KJs

★ Over a larger sample size, this chart will smooth out.

- In this sample, I was only dealt AKs a total of 6 times, so a few bad hands can easily skew the data - especially for a hand like AK that likes to go all in before the flop, which causes extremely high variance
- Additionally, 88 and 77 are basically the same hand in theory, but 88 was a big winner and 77 was a big loser, simply because of the small sample size

Most and Least Profitable Hands - Opponents



AA	AKs	AQs	AJs	ATs	A9s	A8s	A7s	A6s	A5s	A4s	A3s	A2s
AKo	KK	KQs	KJs	KTs	K9s	K8s	K7s	K6s	K5s	K4s	K3s	K2s
AQo	KQo	QQ	QJs	QTs	Q9s	Q8s	Q7s	Q6s	Q5s	Q4s	Q3s	Q2s
AJo	KJo	QJo	JJ	JTs	J9s	J8s	J7s	J6s	J5s	J4s	J3s	J2s
ATo	KTo	QTo	JTo	TT	T9s	T8s	T7s	T6s	T5s	T4s	T3s	T2s
A9o	K9o	Q9o	J9o	T9o	99	98s	97s	96s	95s	94s	93s	92s
A8o	K8o	Q8o	J8o	T8o	98o	88	87s	86s	85s	84s	83s	82s
A7o	K7o	Q7o	J7o	T7o	97o	87o	77	76s	75s	74s	73s	72s
A6o	K6o	Q6o	J6o	T6o	96o	86o	76o	66	65s	64s	63s	62s
A5o	K5o	Q5o	J5o	T5o	95o	85o	75o	65o	55	54s	53s	52s
A4o	K4o	Q4o	J4o	T4o	94o	84o	74o	64o	54o	44	43s	42s
A3o	K3o	Q3o	J3o	T3o	93o	83o	73o	63o	53o	43o	33	32s
A2o	K2o	Q2o	J2o	T2o	92o	82o	72o	62o	52o	42o	32o	22

Green = net profit **Red = net loss**

- Five most profitable hands:
  - **K8s,** AKs, QQ, T7s, QJs
- Five least profitable hands:
  - **Q7s**, A5s, A4s, A6s, T6s
- Outside of AA and AK, opponents lost money on every single hand that contains an ace! Why?
  - Ace/rag hands (an ace with a low second card) actually aren't very good hands, but inexperienced players often don't know this and play them too much.
  - These hands are bad because it is very easy to make top pair (with any ace on the board) but lose because your opponent also has an ace, but with a better kicker. Even though you often have to do it, bad players aren't very good at folding a pair of aces.
  - Limping with these hands (which many players did) is even worse