

“Shuffle Up and Deal!” Sampling without Replacement Using SAS®

Daniel R. Bretheim, Willis Towers Watson

ABSTRACT

Sampling, whether with or without replacement, is an important component of the hypothesis testing process. In this paper, we demonstrate the mechanics and outcome of sampling without replacement, where sample values are not independent. In other words, what we get in the first sample affects what we can get for the second sample, and so on. We use the popular variant of poker known as No Limit Texas Hold'em to illustrate.

INTRODUCTION

In our No Limit Texas Hold'em example, SAS® will deal cards to nine players from a standard fifty-two card deck. The cards are dealt as follows:

- After the deck is shuffled, each player receives two “hole cards”.
- The first round of betting takes place.
- The dealer “burns” the top card and deals the three card “flop” face up in the middle of the table. These three cards, along with the two that will follow, are called “community cards” because they can be used by everyone who is still in the hand.
- Another round of betting takes place.
- The dealer “burns” another card and reveals the fourth community card, called “the turn”.
- Another round of betting occurs.
- The dealer “burns” one last card and reveals the final community card, called “the river”.
- The final round of betting occurs.

In this scenario, which assumes the hand does not terminate before all possible cards are dealt, we will end up with:

- 18 hole cards
- 3 burned cards
- 5 community cards
- 26 cards remaining in the deck

THE ANATOMY OF SAMPLING WITHOUT REPLACEMENT

Given the need to deal a different number of cards at each stage of the hand, we will use a macro based approach with a positional parameter (number_to_deal) to indicate the number of cards to deal. There are many ways to code sampling without replacement functionality in SAS®. See the code on the next page for one method.

```

%macro deal(dsn,number_to_deal);
data &dsn;
  sampsize=&number_to_deal;
  obsleft=totobs;
  do while(sampsize>0);
    pickit+1;
    if ranuni(0) < sampsize/obsleft then do;
      set deck point=pickit nobs=totobs;
      output;
      sampsize=sampsize-1;
    end;
    obsleft=obsleft-1;
  end;
  stop;
run;
%mend deal;

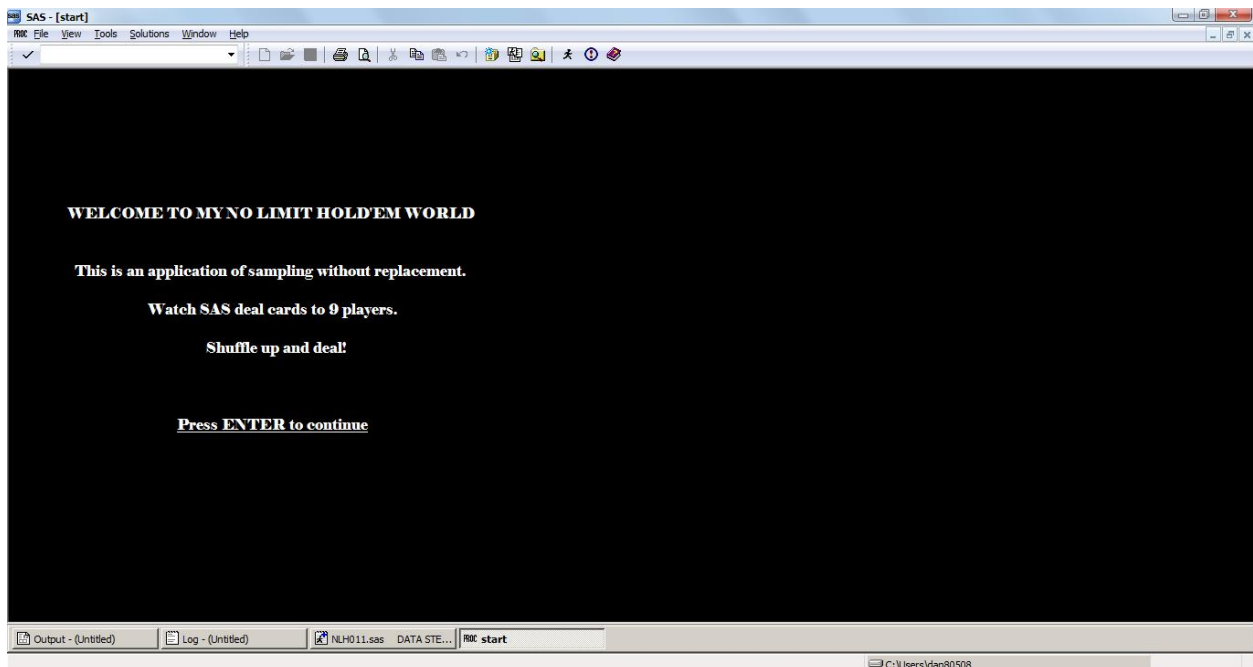
```

Using this code, the POINT= option creates a numeric variable (pickit) whose value is the number of the observation in the input data set (deck) you want the current execution of the SET statement to process. Once a card is selected from the deck (i.e., once each observation is selected from the deck data set), it cannot be reselected for that specific sample, because the value of the pickit variable is incremented each time the IF/THEN/DO loop is executed.

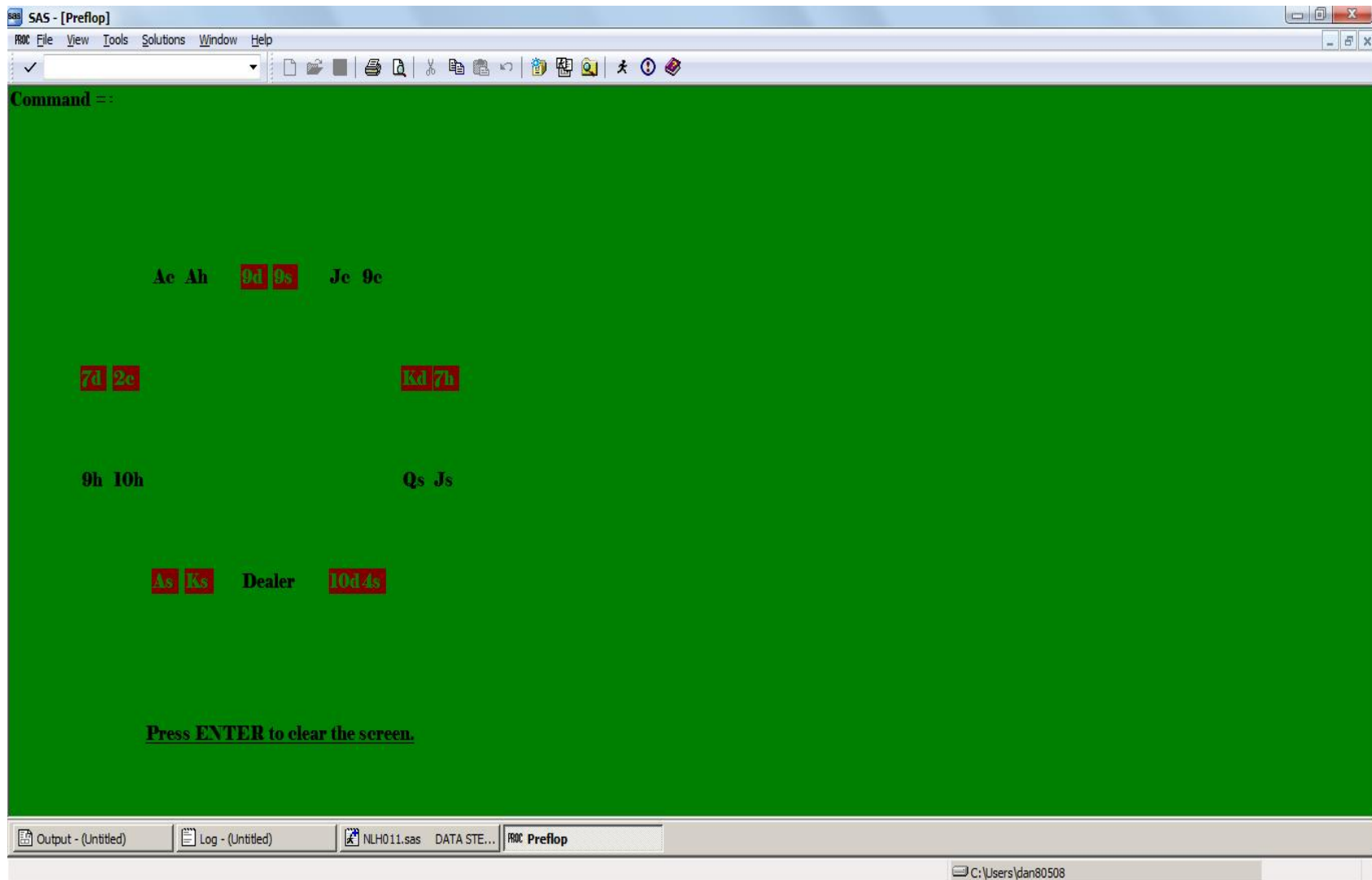
Additional processing utilizes match-merging to eliminate dealt cards from the deck prior to the next dealer action.

DISPLAYING THE RESULTS

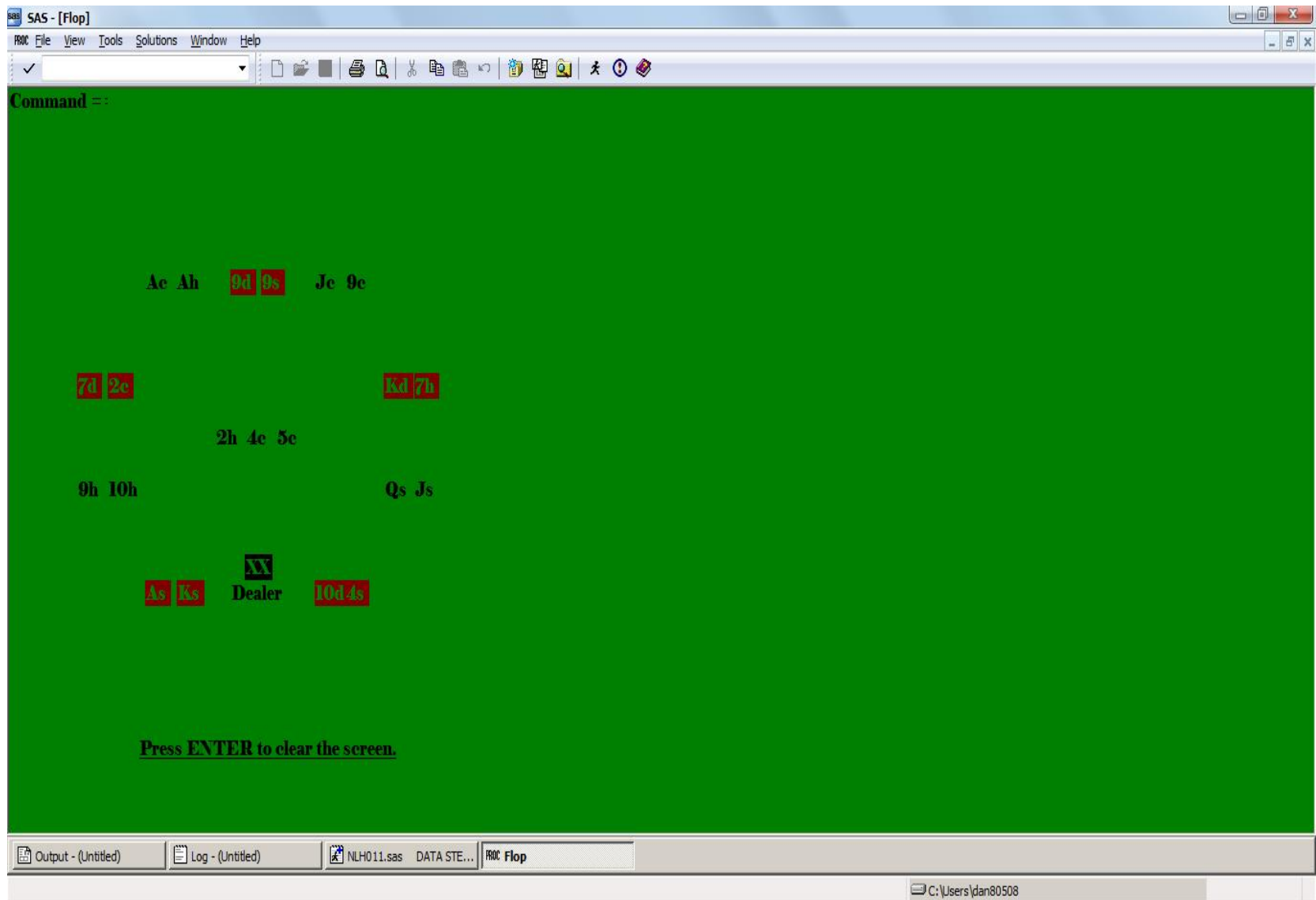
WINDOW and DISPLAY statements are used to display on the screen the cards dealt at each phase of the hand. Here's is an example.



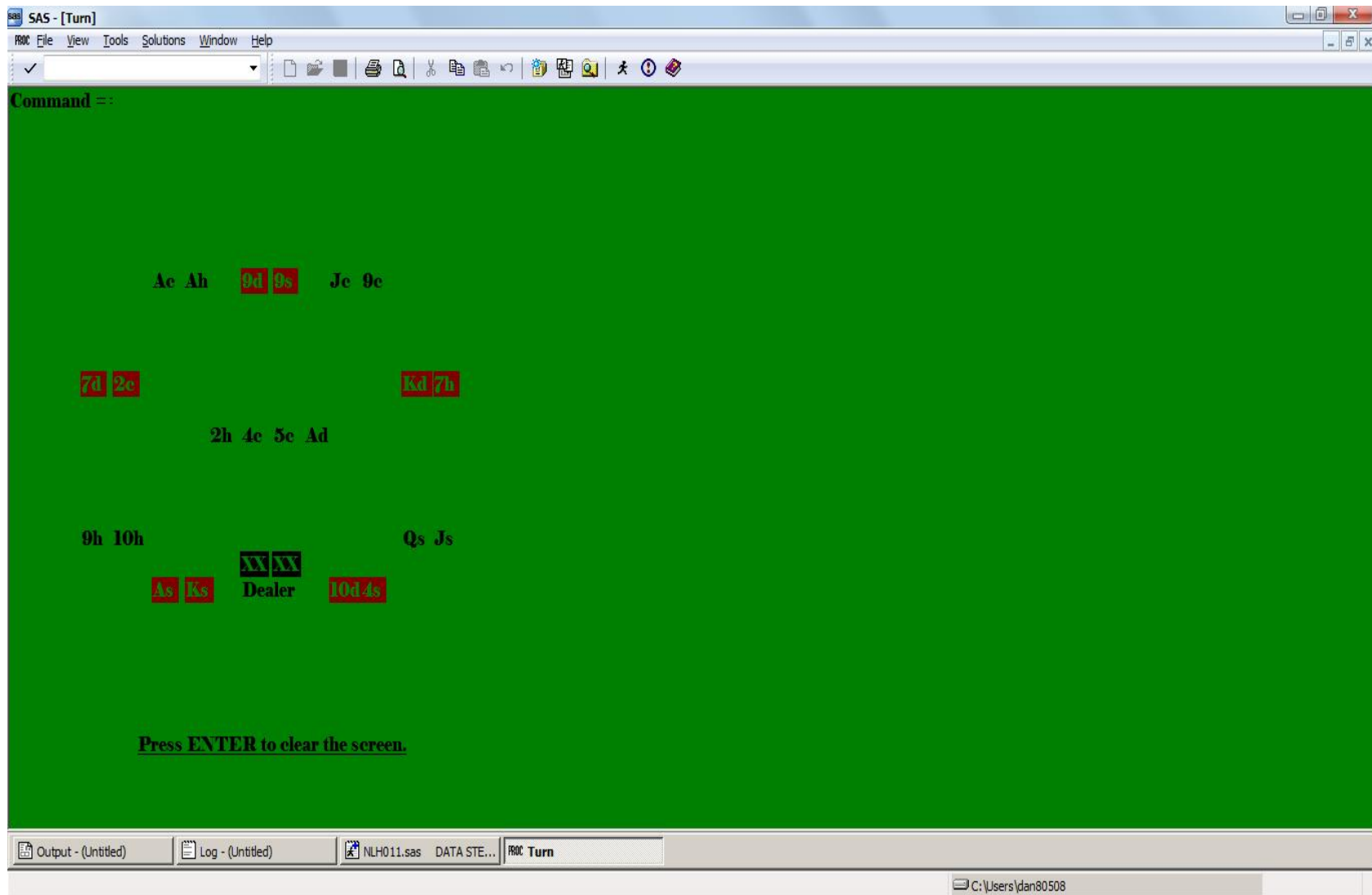
Display 1. Welcome Screen



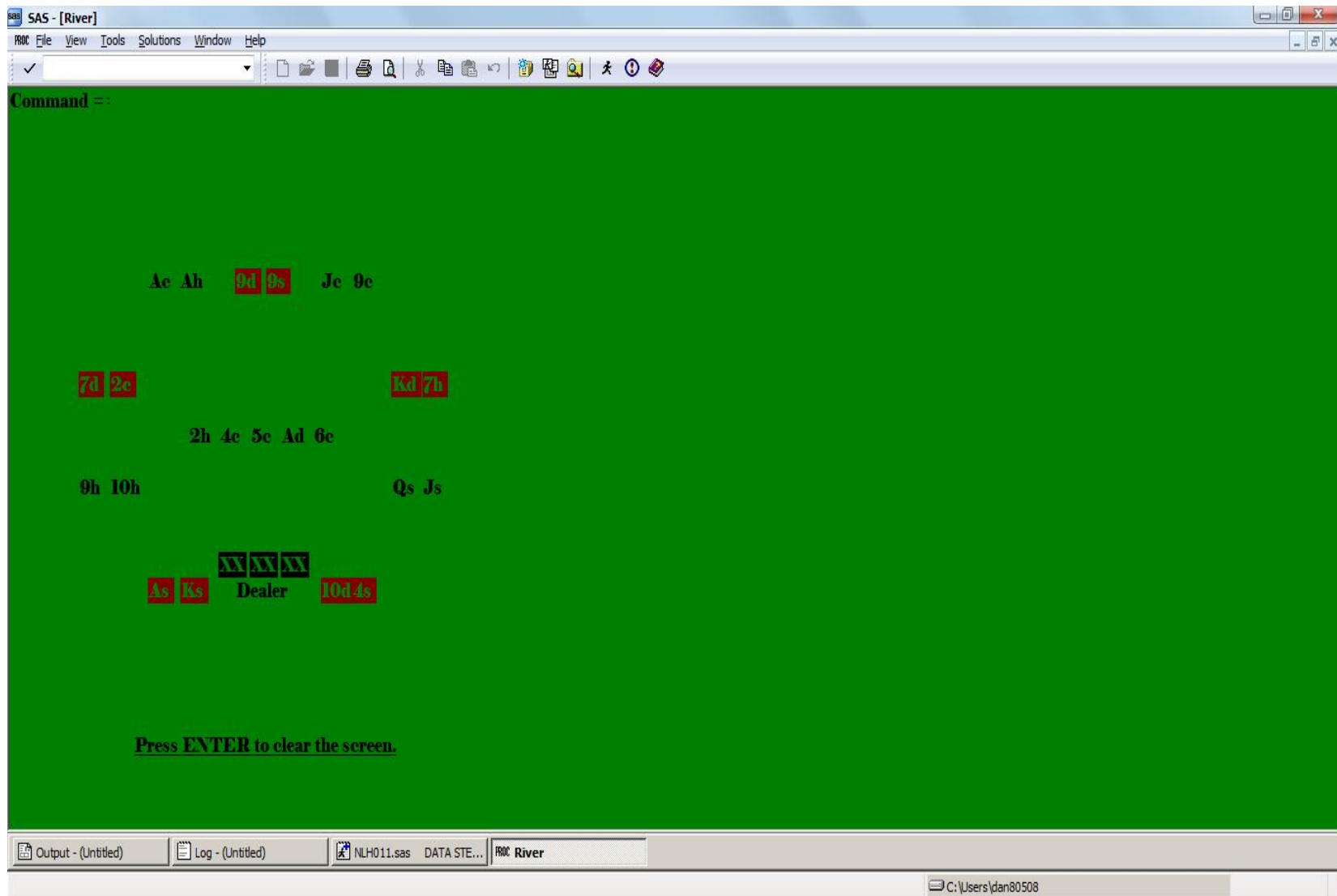
Display 2. Player's Hole Cards



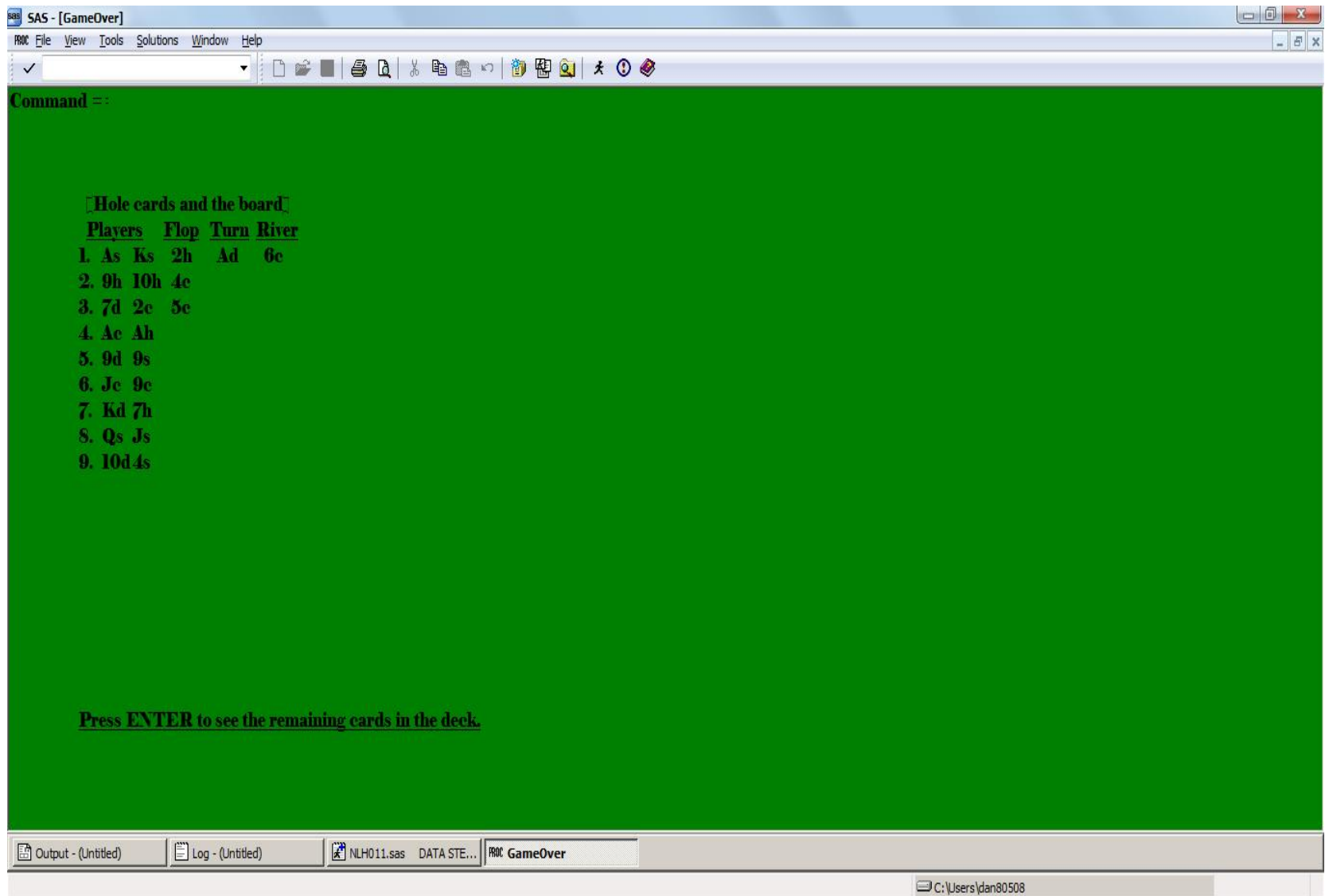
Display 3. The Flop (note the face down burned card above the dealer)



Display 4. The Turn



Display 5. The River



Display 6. Summary of Dealt Cards

SAS - [Remaining]

PROC File View Tools Solutions Window Help

Command ==

Hole cards and the board					Burned	Undealt
Players	Flop	Turn	River	Cards	Cards	Cards
1. As Ks	2h Ad	6c	10c	2s	7c	
2. 9h 10h	4c		10s	3c	7s	
3. 7d 2c	5c		2d	3d	8c	
4. Ac Ah				3h	8d	
5. 9d 9s				3s	8h	
6. Jc 9c				4d	8s	
7. Kd 7h				4h	Jd	
8. Qs Js				5d	Jh	
9. 10d 4s				5h	Kc	
				5s	Kh	
				6d	Qc	
				6h	Qd	
				6s	Qh	

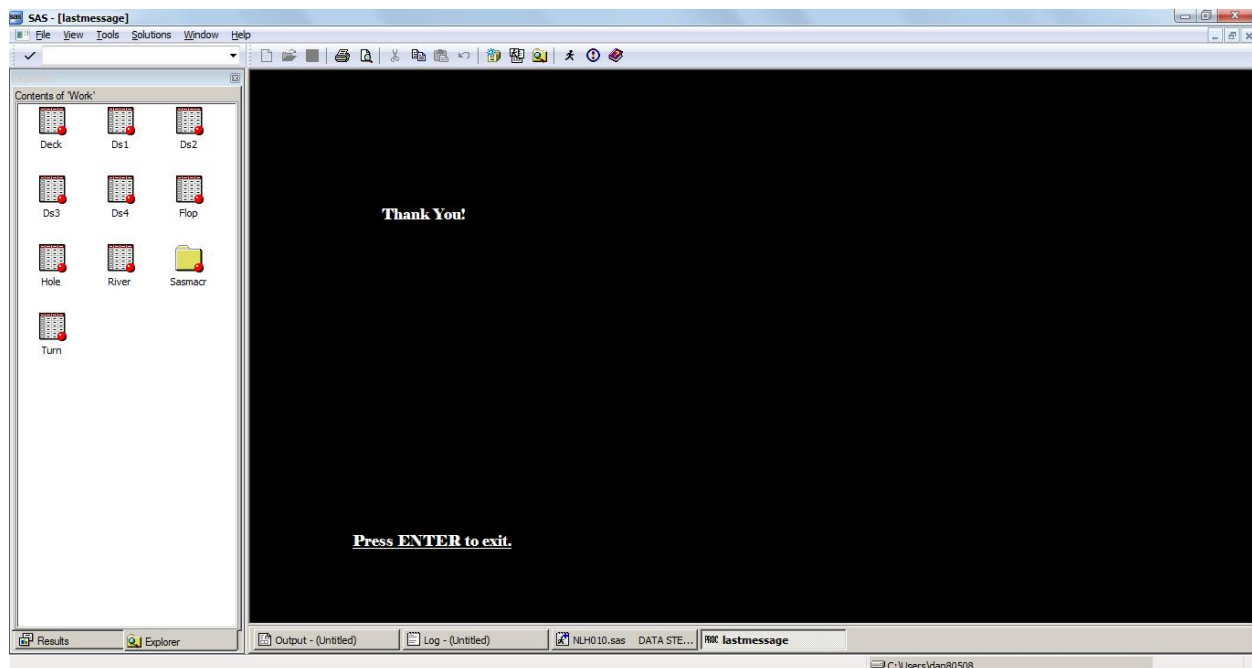
TOTALS=> 18 + 3 + 1 + 1 + 3 + 13 + 13 = 52

Press ENTER to clear the screen.

Output - (Untitled) Log - (Untitled) NLH011.sas DATA STE... PROC Remaining

C:\Users\dan80508

Display 7. All 52 Cards Accounted For



Display 8. Final Screen (note the files used or created when the application runs)

CONCLUSION

Poker is a fun illustration of sampling without replacement, where each observation in the population being sampled can only be selected once. Contact the author at the email address shown below if you would like all the code used in this application.

CONTACT INFORMATION

Contact the author at:

Dan Bretheim
 Willis Towers Watson
 347.278.9638
daniel.bretheim@willistowerswatson.com

SAS and all other SAS Institute Inc. product or service names are registered trademarks or trademarks of SAS Institute Inc. in the USA and other countries. ® indicates USA registration.

Other brand and product names are trademarks of their respective companies.