

SAS® GLOBAL FORUM 2021

Paper 1121-2021

PROC HTTP: Using the Web to Combat Fraud in Auto Finance During a Global Pandemic

Joshua Piers, GM Financial

ABSTRACT

Most car dealerships do not buy cars outright; they finance them until they sell to a consumer. While the vast majority of dealerships conduct themselves respectably, there are occasional cases of dealership fraud. Typically, auditors are sent to dealerships to monitor vehicle inventory and verify cars are on the lot. In the midst of a global pandemic, this has been difficult to do effectively. The practice of selling vehicles without repaying the lender allows a dealership to float cash since it typically only has to pay a monthly interest amount as opposed to the full amount for the vehicle when it sells. Compounded, this practice can set a dealer back beyond the point of return, potentially resulting in hundreds of millions of dollars in loss. Time to get creative! Using PROC HTTP, it's possible to grab HTML code from the dealer inventory web pages, identify VIN numbers from the HTML, and compare the list of VINs to the lender's active loans. As a result, PROC HTTP empowered the lender to monitor dealership inventories in a new way to help overcome the challenges presented by the pandemic.

INTRODUCTION

Pre-pandemic lenders regularly utilized contracted in-person inventory monitoring services on the dealership's physical lots. Teams would go out onto the lot and perform an audit by manually inputting all VINs physically present at the dealership. If a loan existed with the lender for the dealership on a specific VIN, and that VIN was not present at the time of audit, the lender would require confirmation of the VIN's actual physical location.

The pandemic complicated this process. Initially a large number of dealers temporarily shut down operations. Other dealerships limited the amount of people allowed on the premises. There were also instances where dealerships had infections among staff members and were forced to quarantine. All of these factors contributed to not being able to audit dealerships consistently in the traditional sense.

"Selling out of Trust" ("SOT" for short) is a practice where a dealership has a floorplan loan for a vehicle, then sells that vehicle to a consumer, but does not report the vehicle as sold to the lender. Typically, the dealership is required to report the vehicle as sold to the lender within a couple of days of the sale. This practice allows the dealership to continue paying a smaller monthly payment for the vehicle to the lender, while using the dollars provided by the consumer for the vehicle for other purposes (potentially paying bills, buying more inventory, advertising, etc.). Typically, a dealership that engages in an SOT does so because resources are limited. If these conditions at the dealership continue and there are more SOT situations, the dealership can dig itself into a hole it can't escape. This results in large losses for the lender. It is in the best interest of both the dealerships and the lender for consistent accountability practices in inventory monitoring and control.

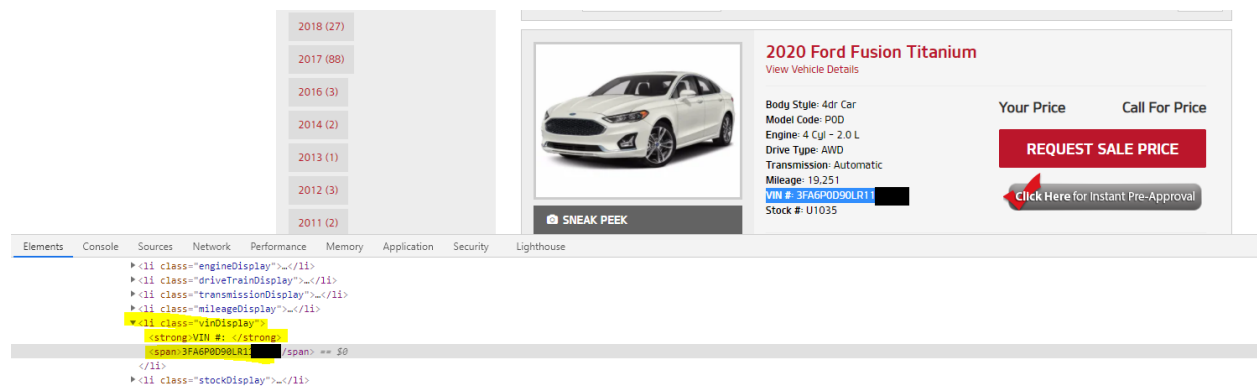
The internet has changed a lot of things in the way dealerships do business. Modern consumers are well informed and arrive at dealerships having done their research online.

Not only that, but a lot of consumers are now opting to cut out the trip to the dealership all together and purchase directly online. These new consumer practices require dealerships to keep an up to date inventory list on their websites at all times in order to stay competitive.

PROC HTTP

The pandemic prompted us as a lender to find new and creative ways to continue monitoring dealership inventories, since couldn't continue to do so in the traditional ways. I was introduced to PROC HTTP at SAS Global Forum 2019 in Dallas at a breakout session. I made a mental note that it seemed like a really cool tool, but at the time I couldn't think of any practical applications for the business. The pandemic changed that. I started exploring whether we could identify VINs from a dealership's advertised website inventory and compare this advertised list to the list of the lender's active floorplan VINs. In every case I researched, VIN data was stored in the html code. In a lot of cases the VIN was advertised directly on the site, but in some cases it existed only on the back end as a unique identifier.

If you are attempting to do your own research, this simplest way I have found to do so is in Chrome. On the main inventory list page you will want to identify a vehicle, right-click and select "Inspect", and then search through all of the html from that section of the page until you find a VIN number. I did run into an instance where the HTML referenced a widget and did not contain a VIN number, but when I navigated to the widget that was referenced, I discovered all of the VIN numbers in the widget data.



Once you confirm that the VIN numbers exist and identify their location, you can start pulling them into SAS using PROC HTTP. In the below image, there are a couple of things to note:

- Most lenders will have security setup that makes it difficult to use PROC HTTP. In my case, I needed to use a service account (below where you see proxyusername and proxypassword). I also needed to submit each specific dealership's website to our cyber security team so the service account could grab the data and pull it into SAS.
- Please note the two different URLs below. Most dealer websites have a maximum display number for inventory pages. I've seen some that go up to 200, and some that only go to 16. The case below allowed 100 per page. You will need to run PROC HTTP for each inventory page until all available pages have been captured. In the case below I manually added them one by one, but you could potentially run a loop where you grab these pages until there's no more data to import.
- Each PROC HTTP statement send the data to a unique txt file. You will also need to create a txt file for each individual inventory page.

- It's also important to note that a lot of dealerships keep their used inventory and new inventory in separate lists. Used inventory is the hardest for a lender to monitor and represents the greatest risk in most cases.

```
filename resp '/SharePoint/[REDACTED]_used_inventory.txt';

proc http
  url="https://www.[REDACTED].com/searchused.aspx?pn=100"
  method="get"
  out=resp
  proxyhost="[REDACTED]"
  proxyport=9191
  proxyusername="[REDACTED]"
  proxypassword="[REDACTED]"
;run;

filename resp '/SharePoint/[REDACTED]_used_inventory2.txt';

proc http
  url="https://www.[REDACTED].com/searchused.aspx?pn=100&pt=2"
  method="get"
  out=resp
  proxyhost="[REDACTED]"
  proxyport=9191
  proxyusername="[REDACTED]"
  proxypassword="[REDACTED]"
;run;
```

Now that you've grabbed your data from the web it's time to import it into SAS. I did not have any prior experience importing text files into SAS before working on this project. The code in the image below was built out using the "Import Data" function in SAS. Since the HTML code varies so much from site to site, and in a lot of cases there are not great delimiters, I found it most helpful to import everything as a single column and cleanup from there. You will need to complete this step for each of the .txt files you created using PROC HTTP. Note the character limit below. You will want to validate that nothing got cut off when you import your code.

```

/* We must now import the files. Column name and character length may need to be adjusted accordingly. */
■ DATA WORK. ██████████_USED_INVENTORY;
  LENGTH
  __DOCTYPE_html_ $ 32767 ;
  LABEL
  __DOCTYPE_html_ = "<!DOCTYPE html>" ;
  FORMAT
  __DOCTYPE_html_ $CHAR32767. ;
  INFORMAT
  __DOCTYPE_html_ $CHAR32767. ;
  INFILE '/SharePoint/██████████_used_inventory.txt'
  LRECL=32767
  FIRSTOBS=2
  ENCODING="LATIN1"
  TRUNCOVER ;
  INPUT
  @1 __DOCTYPE_html_ $CHAR32767. ;
  RUN;

■ DATA WORK. ██████████_used_INVENTORY2;
  LENGTH
  __DOCTYPE_html_ $ 32767 ;
  LABEL
  __DOCTYPE_html_ = "<!DOCTYPE html>" ;
  FORMAT
  __DOCTYPE_html_ $CHAR32767. ;
  INFORMAT
  __DOCTYPE_html_ $CHAR32767. ;
  INFILE '/SharePoint/██████████_used_inventory2.txt'
  LRECL=32767
  FIRSTOBS=2
  ENCODING="LATIN1"
  TRUNCOVER ;
  INPUT
  @1 __DOCTYPE_html_ $CHAR32767. ;
  RUN;

```

Once imported, you should end up with tables that look like this:

	__DOCTYPE_html__
1	</head><body><!-- GoogleTagManager PageBodyTop (noscript) --><noscript><iframe...
2	<h1 class="header">
3	Search Inventory
4	</h1>
5	</div>
6	<div id="content-main" class="Search-Inventory-page"><div class='conditionalBlocks'>...
7	<div class="vehiclePhoto stat-thumbnail" data-loc="available inventory" id='vehiclePho...
8	</div>
9	<div class="col-md-12"><ul class="vehicleImgLinks list-margin-top_5x"><li id="sneakP...
10	<div class="vehiclePhoto stat-thumbnail" data-loc="available inventory" id='vehiclePho...
11	</div>
12	<div class="col-md-12 col-sm-12 hidden-xs"><h2 class="vehicleTitle margin-x"><a cla...
13	<div class="vehiclePhoto stat-thumbnail" data-loc="available inventory" id='vehiclePho...
14	</div>
15	<div class="col-md-12"><ul class="vehicleImgLinks list-margin-top_5x"><li id="sneakP...
16	<div class="vehiclePhoto stat-thumbnail" data-loc="available inventory" id='vehiclePho...
17	</div>
18	<div class="col-md-12 col-sm-12 hidden-xs"><h2 class="vehicleTitle margin-x"><a cla...
19	<div class="vehiclePhoto stat-thumbnail" data-loc="available inventory" id='vehiclePho...
20	</div>
21	<div class="col-md-12"><ul class="vehicleImgLinks list-margin-top_5x"><li id="sneakP...
22	<div class="vehiclePhoto stat-thumbnail" data-loc="available inventory" id='vehiclePho...
23	</div>
24	<div class="col-md-12 col-sm-12 hidden-xs"><h2 class="vehicleTitle margin-x"><a cla...
25	<div class="vehiclePhoto stat-thumbnail" data-loc="available inventory" id='vehiclePho...
26	</div>
27	<div class="col-md-12"><ul class="vehicleImgLinks list-margin-top_5x"><li id="sneakP...
28	<div class="vehiclePhoto stat-thumbnail" data-loc="available inventory" id='vehiclePho...
29	</div>
30	<div class="col-md-12 col-sm-12 hidden-xs"><h2 class="vehicleTitle margin-x"><a cla...
31	<div class="vehiclePhoto stat-thumbnail" data-loc="available inventory" id='vehiclePho...
32	</div>

From here you will want to start cleaning up your data. The first thing to do is identify where your VINs are stored and eliminate any non-VIN values. In the image below I do this using the CONTAINS function in PROC SQL and add "USED" as an identifier for inventory type. Once you do this for each of the tables you created in the previous step, you can use PROC APPEND to append them all together.

```

PROC SQL;
CREATE TABLE [REDACTED]_INVENTORY AS SELECT DISTINCT
  "__DOCTYPE_html_"n AS ROUGH_VIN,
  "Used" AS TYPE
FROM WORK.[REDACTED]_USED_INVENTORY
WHERE "__DOCTYPE_html_"n CONTAINS ('<li class="vinDisplay"><strong>VIN #: </strong><span>')
;RUN;

```

Once you append everything together, you should end up with something like this:

	ROUGH_VIN	TYPE
1	<div class="col-md-12 col-sm-12 hidden-xs"><h2 class="vehicleTitle margin-x"><a class="stat-text-link" data-loc="available inventory:1C4RJFBG3HC8	Used
2	<div class="col-md-12 col-sm-12 hidden-xs"><h2 class="vehicleTitle margin-x"><a class="stat-text-link" data-loc="available inventory:1GKKXLSXHZ3	Used
3	<div class="col-md-12 col-sm-12 hidden-xs"><h2 class="vehicleTitle margin-x"><a class="stat-text-link" data-loc="available inventory:1GYKNBRS9H22	Used
4	<div class="col-md-12 col-sm-12 hidden-xs"><h2 class="vehicleTitle margin-x"><a class="stat-text-link" data-loc="available inventory:1GYS4JKJ3HR21	Used
5	<div class="col-md-12 col-sm-12 hidden-xs"><h2 class="vehicleTitle margin-x"><a class="stat-text-link" data-loc="available inventory:1N4AL3AP6FC13	Used
6	<div class="col-md-12 col-sm-12 hidden-xs"><h2 class="vehicleTitle margin-x"><a class="stat-text-link" data-loc="available inventory:2HGFC2F77HH5	Used
7	<div class="col-md-12 col-sm-12 hidden-xs"><h2 class="vehicleTitle margin-x"><a class="stat-text-link" data-loc="available inventory:2T1BURHE3JC01	Used
8	<div class="col-md-12 col-sm-12 hidden-xs"><h2 class="vehicleTitle margin-x"><a class="stat-text-link" data-loc="available inventory:2T2BZMCA3JC14	Used
9	<div class="col-md-12 col-sm-12 hidden-xs"><h2 class="vehicleTitle margin-x"><a class="stat-text-link" data-loc="available inventory:2T2BZMCA7HC0	Used
10	<div class="col-md-12 col-sm-12 hidden-xs"><h2 class="vehicleTitle margin-x"><a class="stat-text-link" data-loc="available inventory:2T2BZMCA9HC0	Used
11	<div class="col-md-12 col-sm-12 hidden-xs"><h2 class="vehicleTitle margin-x"><a class="stat-text-link" data-loc="available inventory:2T3RFREVXH6	Used
12	<div class="col-md-12 col-sm-12 hidden-xs"><h2 class="vehicleTitle margin-x"><a class="stat-text-link" data-loc="available inventory:3C4NJDBB4HT6	Used
13	<div class="col-md-12 col-sm-12 hidden-xs"><h2 class="vehicleTitle margin-x"><a class="stat-text-link" data-loc="available inventory:3FA6P0D90LR11	Used
14	<div class="col-md-12 col-sm-12 hidden-xs"><h2 class="vehicleTitle margin-x"><a class="stat-text-link" data-loc="available inventory:3KPFK4A71JE23	Used
15	<div class="col-md-12 col-sm-12 hidden-xs"><h2 class="vehicleTitle margin-x"><a class="stat-text-link" data-loc="available inventory:3KPFK4A75HE08	Used
16	<div class="col-md-12 col-sm-12 hidden-xs"><h2 class="vehicleTitle margin-x"><a class="stat-text-link" data-loc="available inventory:3N1AB7AP6HY24	Used
17	<div class="col-md-12 col-sm-12 hidden-xs"><h2 class="vehicleTitle margin-x"><a class="stat-text-link" data-loc="available inventory:4JGDA5HB8HA8	Used
18	<div class="col-md-12 col-sm-12 hidden-xs"><h2 class="vehicleTitle margin-x"><a class="stat-text-link" data-loc="available inventory:4JGDA5HBXHA8	Used

From here I used a combination of the PRXCHANGE and the STRIP/TRANWRD functions in a data step to isolate my VIN numbers. If you go this route pay attention to the character length I set at the beginning of the data step. You will need to adjust these as needed as well as the specified language in quotes below as needed. Another great option that's not pictured is using a function like PRXNEXT to try to identify 16-character strings.

```

DATA VIN_LIST;
LENGTH CLEANUP $38;
LENGTH VININVENTORY $18;
SET [REDACTED]_INVENTORY;

CLEANUP = prxchange('s/.*(VIN.*)/$1/', 1, ROUGH_VIN);

VININVENTORY = STRIP(TRANWRD(CLEANUP, 'VIN #: </strong><span>', ''));

ON_WEB = 1;

run;

```

The result leaves you with a clean VIN number:

	CLEANUP	VINVENTORY	ROUGH_VIN	TYPE	ON_WEB
1	VIN #. 1C4RJFBG3HC6	1C4RJFBG3HC6	<div class="col-md-12 col-sm-12 hidden-xs"><h2 class="vehicleTitle margin-x"><a cla...	Used	1
2	VIN #. 1GKKNXLSXHZ3	1GKKNXLSXHZ3	<div class="col-md-12 col-sm-12 hidden-xs"><h2 class="vehicleTitle margin-x"><a cla...	Used	1
3	VIN #. 1GYKNBRS9HZ2	1GYKNBRS9HZ2	<div class="col-md-12 col-sm-12 hidden-xs"><h2 class="vehicleTitle margin-x"><a cla...	Used	1
4	VIN #. 1GYS4JKJ3HR2	1GYS4JKJ3HR2	<div class="col-md-12 col-sm-12 hidden-xs"><h2 class="vehicleTitle margin-x"><a cla...	Used	1
5	VIN #. 1N4AL3AP6FC1	1N4AL3AP6FC1	<div class="col-md-12 col-sm-12 hidden-xs"><h2 class="vehicleTitle margin-x"><a cla...	Used	1
6	VIN #. 2HGFC2F77HH5	2HGFC2F77HH5	<div class="col-md-12 col-sm-12 hidden-xs"><h2 class="vehicleTitle margin-x"><a cla...	Used	1
7	VIN #. 2T1BURHE3JC0	2T1BURHE3JC0	<div class="col-md-12 col-sm-12 hidden-xs"><h2 class="vehicleTitle margin-x"><a cla...	Used	1
8	VIN #. 2T2BZMCA3JC1	2T2BZMCA3JC1	<div class="col-md-12 col-sm-12 hidden-xs"><h2 class="vehicleTitle margin-x"><a cla...	Used	1
9	VIN #. 2T2BZMCA7HC0	2T2BZMCA7HC0	<div class="col-md-12 col-sm-12 hidden-xs"><h2 class="vehicleTitle margin-x"><a cla...	Used	1
10	VIN #. 2T2BZMCA9HC0	2T2BZMCA9HC0	<div class="col-md-12 col-sm-12 hidden-xs"><h2 class="vehicleTitle margin-x"><a cla...	Used	1
11	VIN #. 2T3RFREXHW	2T3RFREXHW	<div class="col-md-12 col-sm-12 hidden-xs"><h2 class="vehicleTitle margin-x"><a cla...	Used	1
12	VIN #. 3C4NJDD84HT6	3C4NJDD84HT6	<div class="col-md-12 col-sm-12 hidden-xs"><h2 class="vehicleTitle margin-x"><a cla...	Used	1
13	VIN #. 3FACFP09L011	3FACFP09L011	<div class="col-md-12 col-sm-12 hidden-xs"><h2 class="vehicleTitle margin-x"><a cla...	Used	1

Once you have a clean VIN number you can take the VIN list you created from the dealer's website and compare it to a list of the lender's active floorplan loans. Any VINs on the active floorplan list that are not found advertised on the dealership's website provides the lender the ability to ask specific questions about specific vehicles. Most dealerships will keep their website inventory up to date for the sake of their customers. There is a bit of work on the front end to get everything setup for each dealer, but (barring any major website overhauls) the lender can run this code daily to monitor any changes in dealership inventory.

CONCLUSION

PROC HTTP provides invaluable insight for lenders to dealership inventories and helps overcome obstacles presented to traditional inventory motoring and control in the auto finance industry.

CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the author at:

Joshua Piers
 GM Financial – Commercial Lending
 214-210-5486
Joshua.Piers@GMFinancial.com