



Final Project - Process Mining

Course: Business Process Management
1400-1401

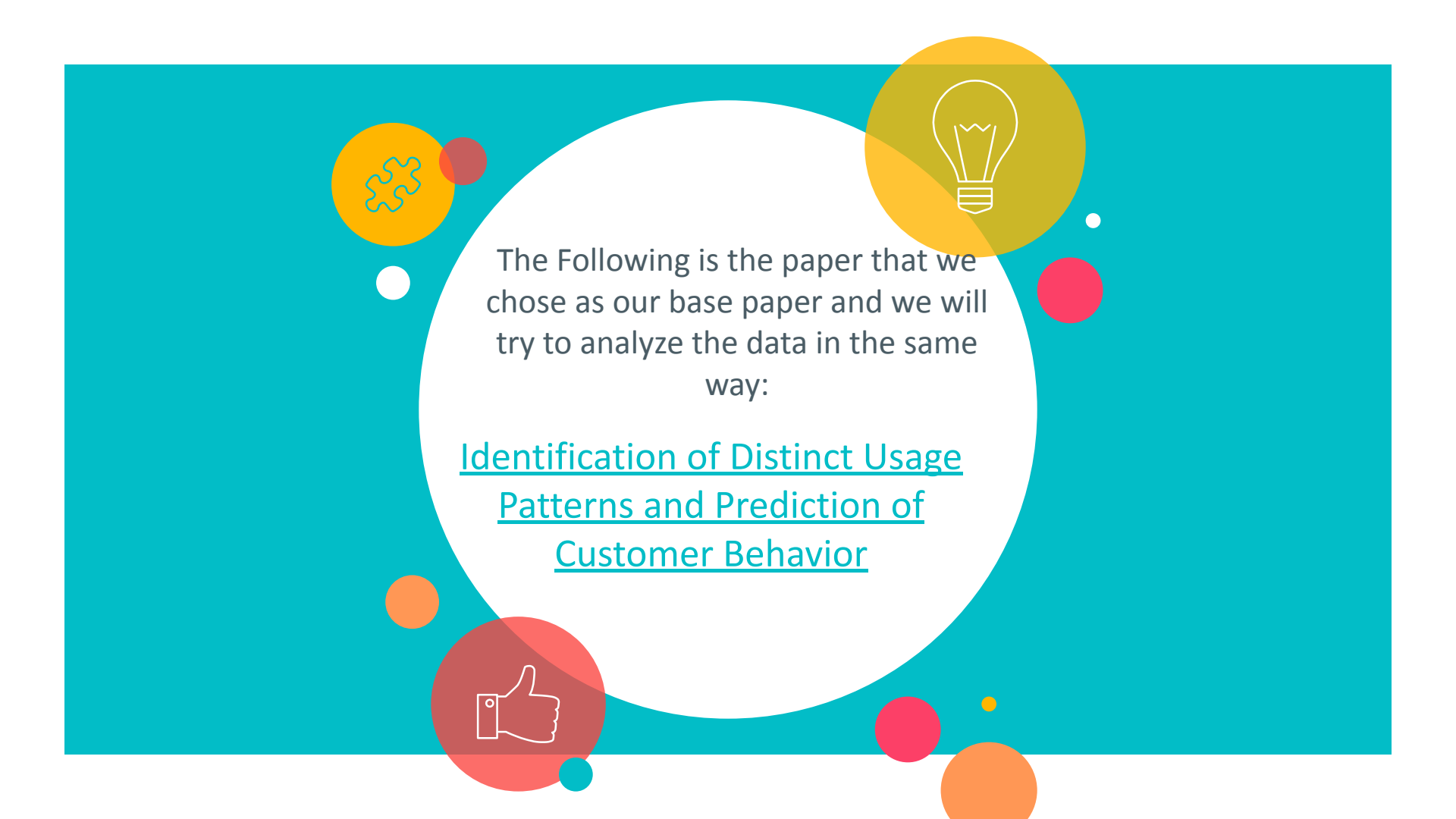
Professor: Dr. Erfan Hassannayebi



Sina
Aghaee

Farzaneh
Mahmoudi

Sahar
Varchandi



The Following is the paper that we chose as our base paper and we will try to analyze the data in the same way:

Identification of Distinct Usage
Patterns and Prediction of
Customer Behavior




About the Dataset and the Paper

Our data belongs to UWV and presented in BPI 2016 Challenge.

UWV (Employee Insurance Agency) is a Dutch autonomous administrative authority (ZBO) and is commissioned by the Ministry of Social Affairs and Employment (SZW) to implement employee insurances and provide labour market and data services in the Netherlands.

The Dutch employee insurances are provided for via laws such as the WW (Unemployment Insurance Act), the WIA (Work and Income according to Labour Capacity Act, which contains the IVA (Full Invalidation Benefit Regulations), WGA (Return to Work (Partially Disabled) Regulations), the Wajong (Disability Assistance Act for Handicapped Young Persons), the WAO (Invalidity Insurance Act), the WAZ (Self-employed Persons Disability Benefits Act), the Wvz (Work and Care Act) and the Sickness Benefits Act.





Data has been collected from several different sources, namely:



1) Click data from the site www.werk.nl collected from visitors that were not logged in:

[BPI Challenge 2016: Clicks NOT Logged In](#)

2) Click data from the customer specific part of the site www.werk.nl (a link is made with the customer that logged in):

[BPI Challenge 2016: Clicks Logged In](#)

3) Werkmap Message data, showing when customers contacted the UWV through a digital channel:

[BPI Challenge 2016: Werkmap Messages](#)

4) Call data from the call center, showing when customers contacted the call center by phone:

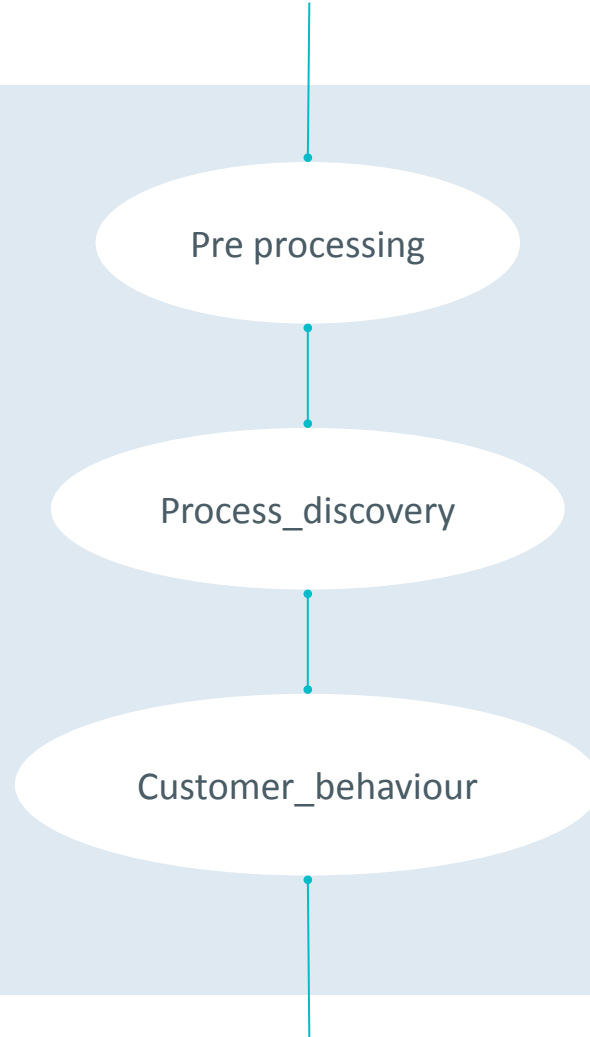
[BPI Challenge 2016: Questions](#)

5) Complaint data showing when customers complained:

[BPI Challenge 2016: Complaints](#)



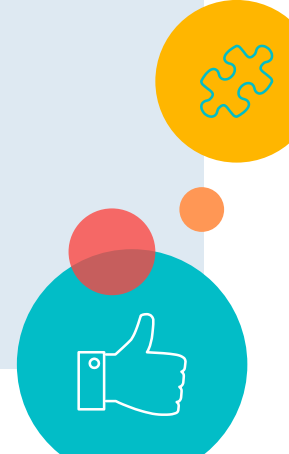
Our process :





Clicks logged in
Selected Columns

We only need these columns for our analysis and process discovery:

- * CustomerID
 - * SessionID
 - * AgeCategory
 - * Gender
 - * TIMESTAMP
 - * PAGE_NAME
- 



7,174,934

There is about 7 milloin Records in clicks_logged_in dataset!!!

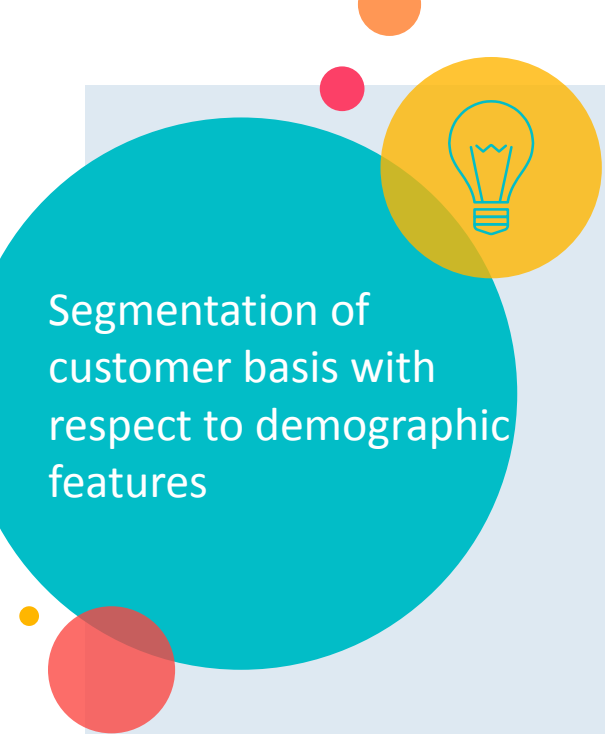




600


WOW!!! We have 600 activities!!!! That's too much! We will probably work with the most frequent ones, and since our activities are the web pages that users visited, it makes sense!





Segmentation of customer basis with respect to demographic features

we segmented our data in the same way our chosen article has done (the article's tables only show the result for the first four); based on the demographic information. We will segment our data into six different data sets. We export all in CSV format for further analysis (Visualization with PowerBI) :

- * Segment 1: Age 18-29
 - * Segment 2: Age 30-39
 - * Segment 3: Age 40-49
 - * Segment 4: Age 50-65
 - * Segment 5: Females
 - * Segment 6: Males
- 

	Segment	Number Of Sessions	Number Of Customers	Number of Events
0	Age category 18-29	105,870	7,092	1,106,045
1	Age category 30-39	133,356	6,420	1,432,107
2	Age category 40-49	158,732	5,891	1,780,253
3	Age category 50-65	262,548	7,244	2,856,529
4	Total	660,506	26,647	7,174,934

Table from our base Article

Table 3. Segmentation of customer basis with respect to demographic features.

Segment	Cases	Events
#1 Age category 18-29	105.832	1.102.717
#2 Age category 30-39	133.310	1.427.763
#3 Age category 40-49	158.687	1.774.864
#4 Age category 50-65	262.490	2.848.044

Our Table from python codes

Activities frequency for all the logged_in dataset



In the table, you can see the most frequent activities for all logged-in customers (those with more than one percent relative frequency), as you see only 14 out of 600 hundred webpages visited in more than one percent of the time.

	Activity	Absolute Frequency	Relative Frequency(%)
0	taken	1,823,175	25.41
1	vacatures_bij_mijn_cv	953,969	13.30
2	mijn_cv	880,597	12.27
3	home	583,545	8.13
4	vacatures_zoeken	582,645	8.12
5	mijn_berichten	529,311	7.38
6	aanvragen-ww	251,063	3.50
7	mijn_werkmap	207,776	2.90
8	mijn_sollicitaties	203,833	2.84
9	werkmap	181,865	2.53
10	inschrijven	148,872	2.07
11	mijn_documenten	143,105	1.99
12	vacatures	78,837	1.10
13	vragenlijst-uwv	74,604	1.04





Activities frequency for the segment 1

Our Table from python codes

	Activity	Absolute Frequency	Relative Frequency(%)
0	taken	327,260	29.59
1	mijn_cv	167,479	15.14
2	home	84,030	7.60
3	vacatures_bij_mijn_cv	82,818	7.49
4	mijn_berichten	79,695	7.21
5	vacatures_zoeken	67,415	6.10
6	aanvragen-ww	59,514	5.38
7	inschrijven	36,224	3.28
8	mijn_werkmap	29,060	2.63
9	mijn_sollicitaties	27,353	2.47
10	mijn_documenten	21,300	1.93
11	werkmap	18,941	1.71
12	wijziging_doorgeven	14,058	1.27
13	vragenlijst-uwv	13,100	1.18

Table 4. Filtered activities representing the most frequent activities for segment 1.

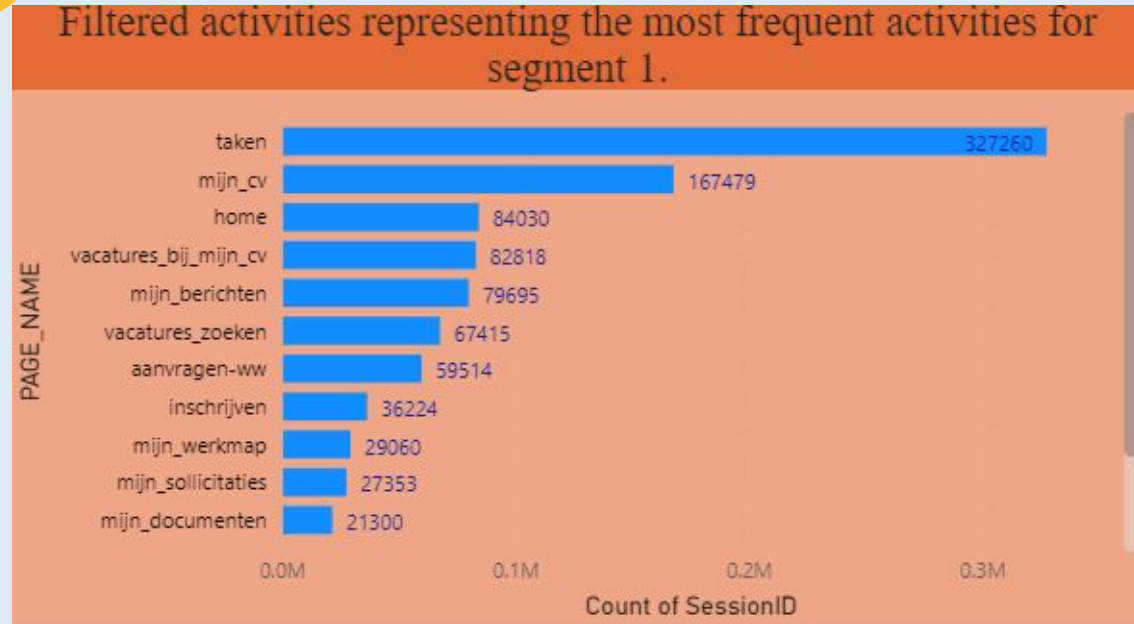
Activity	Absolute Frequency	Relative Frequency
taken	326.294	29,59%
mijn_cv	166.920	15,14%
home	83.778	7,60%
vacatures_bij_mijn_cv	82.588	7,49%
mijn_berichten	79.481	7,21%
vacatures_zoeken	67.217	6,10%
aanvragen-ww	59.339	5,38%
inschrijven	36.113	3,27%
mijn_werkmap	28.967	2,63%
mijn_sollicitaties	27.268	2,47%
mijn_documenten	21.237	1,93%
werkmap	18.882	1,71%
wijziging_doorgeven	14.026	1,27%
vragenlijst-uwv	13.058	1,18%

Our Table from python codes

As you see in the above table, the absolute frequencies have a slight difference with the following table. We mentioned this before that the reason is the authors deleted some rows which we dont know why!! However the relative frequency we calculated is the same as the numbers in the articles table.



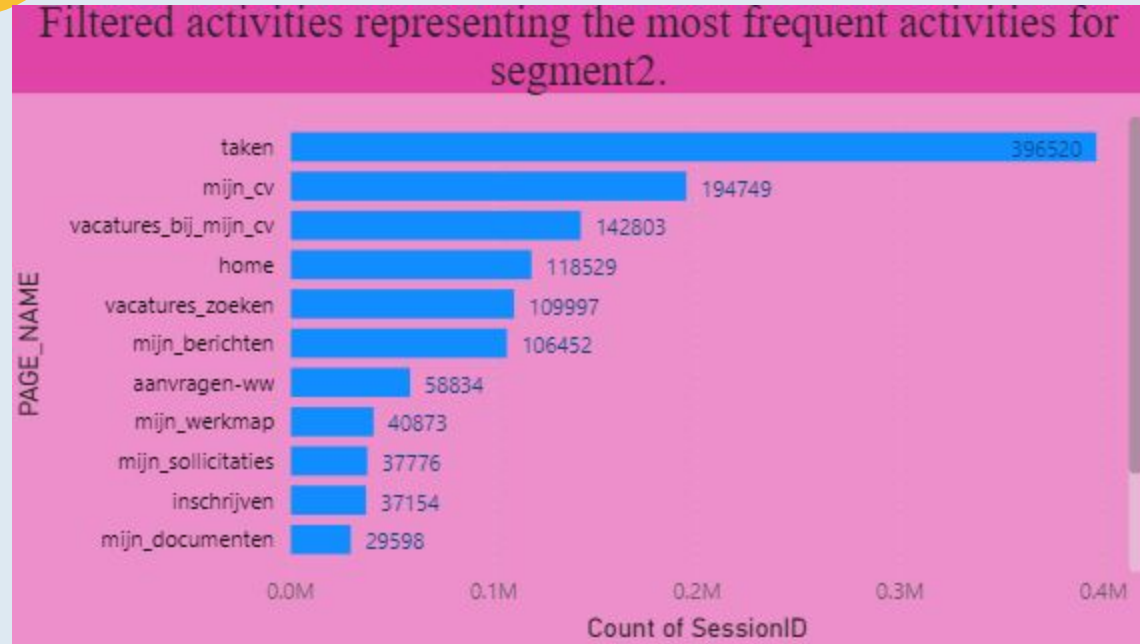
The following chart created by PowerBI on the same dataset in seperate analysis:



Activities
frequency for the
segment 2 :

	Activity	Absolute Frequency	Relative Frequency(%)
0	taken	396,520	27.69
1	mijn_cv	194,749	13.60
2	vacatures_bij_mijn_cv	142,803	9.97
3	home	118,529	8.28
4	vacatures_zoeken	109,997	7.68
5	mijn_berichten	106,452	7.43
6	aanvragen-ww	58,834	4.11
7	mijn_werkmap	40,873	2.85
8	mijn_sollicitaties	37,776	2.64
9	inschrijven	37,154	2.59
10	mijn_documenten	29,598	2.07
11	werkmap	27,487	1.92
12	vragenlijst-uwv	16,144	1.13
13	vacatures	14,558	1.02

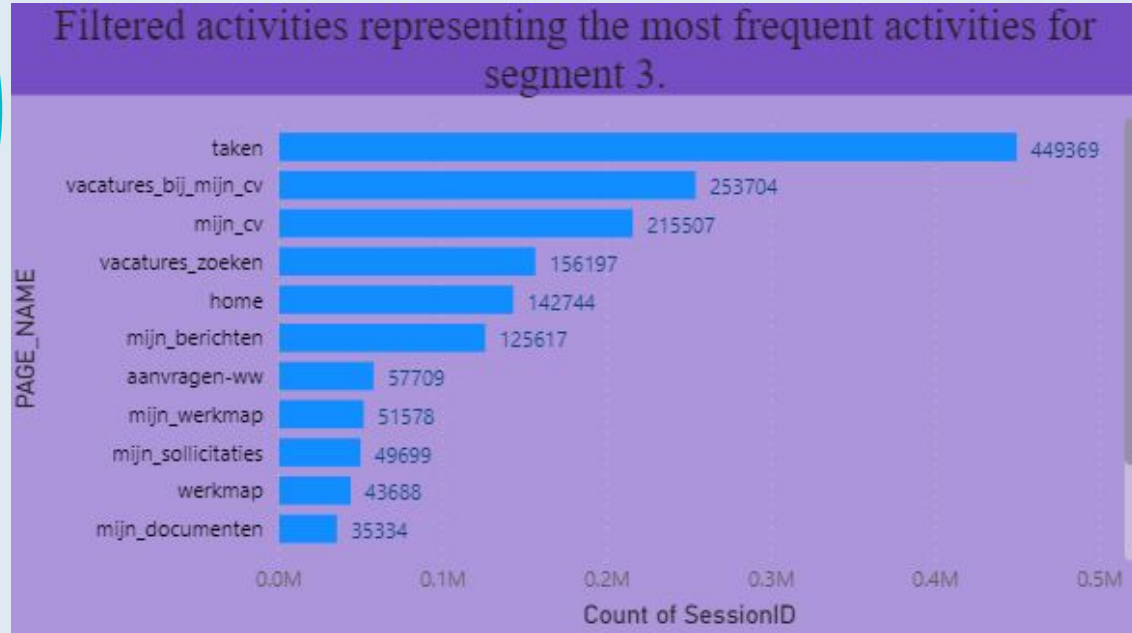
The following chart created by PowerBI on the same dataset in separate analysis:



Activities
frequency for the
segment 3 :

	Activity	Absolute Frequency	Relative Frequency(%)
0	taken	449,369	25.24
1	vacatures_bij_mijn_cv	253,704	14.25
2	mijn_cv	215,507	12.11
3	vacatures_zoeken	156,197	8.77
4	home	142,744	8.02
5	mijn_berichten	125,617	7.06
6	aanvragen-ww	57,709	3.24
7	mijn_werkmap	51,578	2.90
8	mijn_sollicitaties	49,699	2.79
9	werkmap	43,688	2.45
10	mijn_documenten	35,334	1.98
11	inschrijven	33,385	1.88
12	vacatures	21,081	1.18
13	vragenlijst-uwv	18,920	1.06

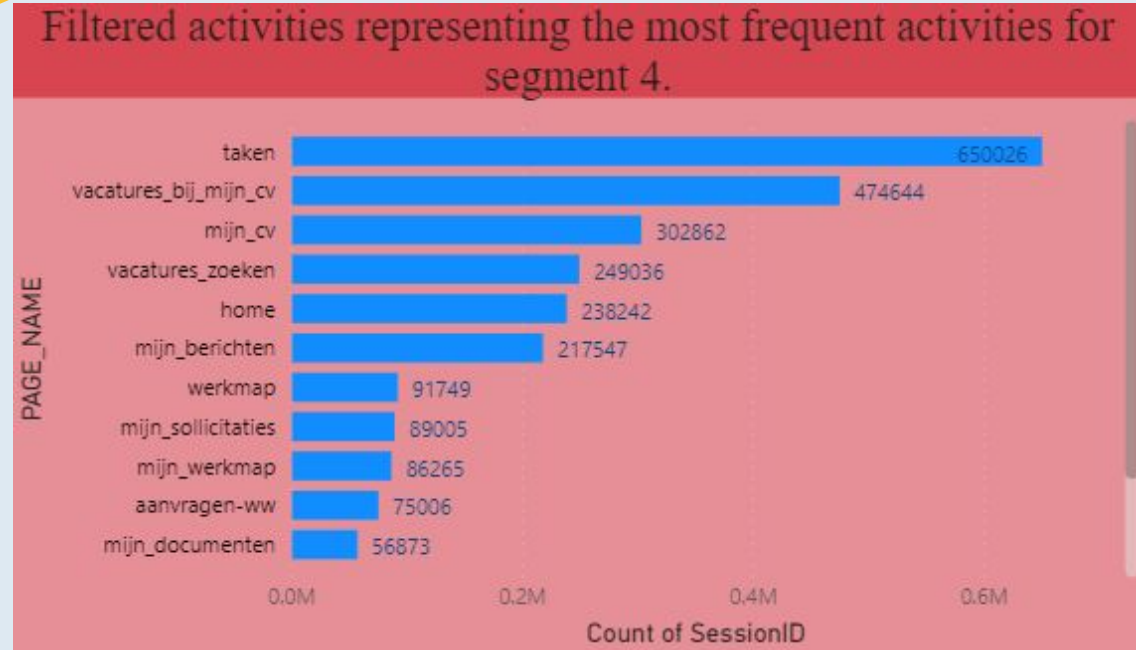
The following chart created by PowerBI on the same dataset in separate analysis:



Activities
frequency for the
segment 4 :

	Activity	Absolute Frequency	Relative Frequency(%)
0	taken	650,026	22.76
1	vacatures_bij_mijn_cv	474,644	16.62
2	mijn_cv	302,862	10.60
3	vacatures_zoeken	249,036	8.72
4	home	238,242	8.34
5	mijn_berichten	217,547	7.62
6	werkmap	91,749	3.21
7	mijn_sollicitaties	89,005	3.12
8	mijn_werkmap	86,265	3.02
9	aanvragen-ww	75,006	2.63
10	mijn_documenten	56,873	1.99
11	inschrijven	42,109	1.47
12	vacatures	32,285	1.13

The following chart created by PowerBI on the same dataset in separate analysis:





Activities frequency for the segment 6

	Activity	Absolute Frequency	Relative Frequency(%)
0	taken	853,416	23.04
1	vacatures_bij_mijn_cv	539,185	14.56
2	mijn_cv	424,378	11.46
3	home	282,645	7.63
4	mijn_berichten	249,794	6.74
5	vacatures_zoeken	237,272	6.41
6	aanvragen-ww	131,524	3.55
7	mijn_werkmap	109,638	2.96
8	mijn_sollicitaties	95,832	2.59
9	werkmap	84,529	2.28
10	inschrijven	77,580	2.09
11	mijn_documenten	65,536	1.77

Activities frequency for the segment 5

20

	Activity	Absolute Frequency	Relative Frequency(%)
0	taken	969,759	26.18
1	mijn_cv	456,219	12.32
2	vacatures_bij_mijn_cv	414,784	11.20
3	vacatures_zoeken	345,373	9.32
4	home	300,900	8.12
5	mijn_berichten	279,517	7.55
6	aanvragen-ww	119,539	3.23
7	mijn_sollicitaties	108,001	2.92
8	mijn_werkmap	98,138	2.65
9	werkmap	97,336	2.63
10	mijn_documenten	77,569	2.09
11	inschrijven	71,292	1.92
12	vacatures	43,702	1.18
13	vragenlijst-uwv	39,375	1.06





Challenge 1: Distinct Usage Patterns



Steps:

1. Most Frequent Activities(Relative Frequency ≥ 1)
2. Selecting Columns (Session ID, Page Name, Timestamp)
3. Making a Dataframe
4. Transforming to XES
5. Heuristic Miner

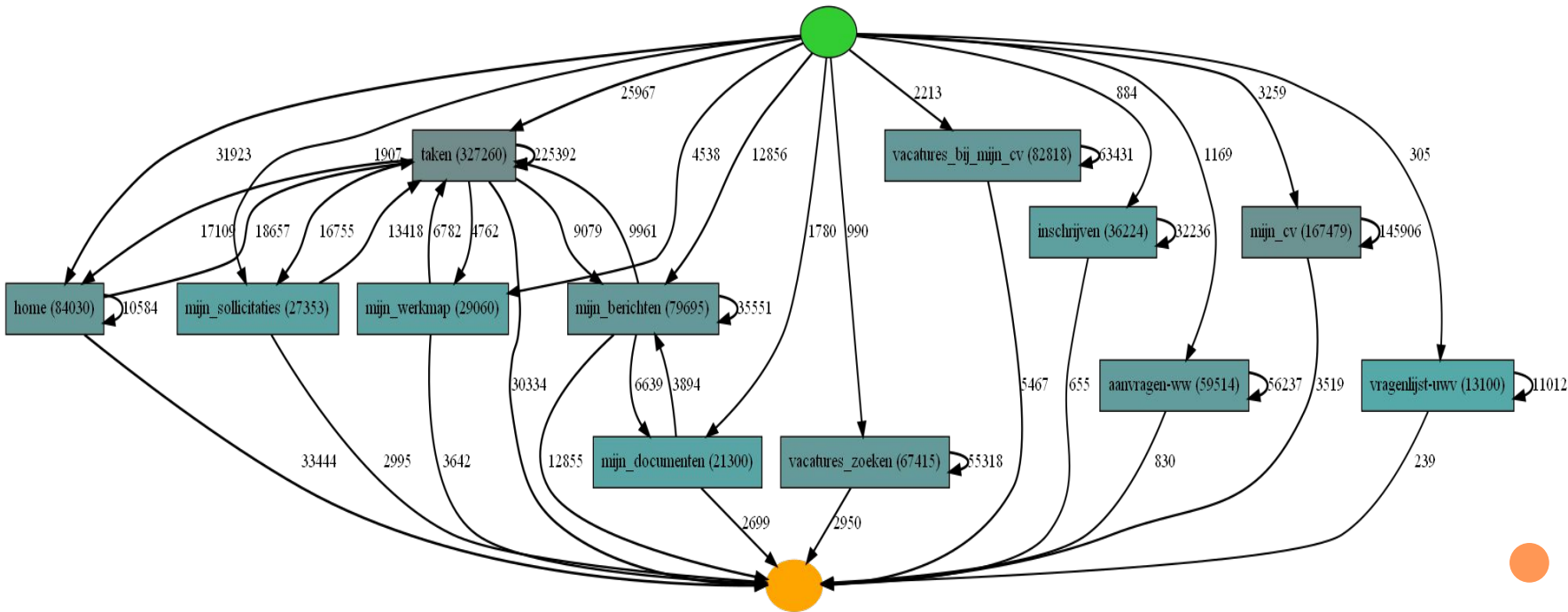
```
net(  
  hold=0.9999,
```

```
t, file_path='Segme
```

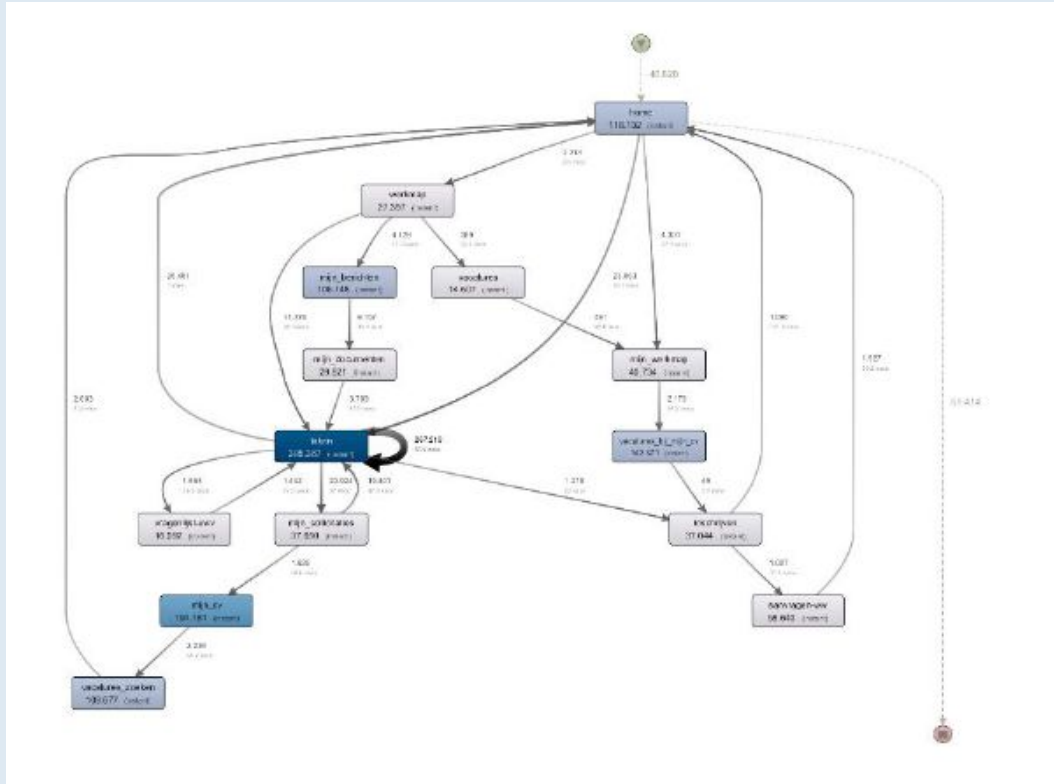


Segment 1: Age 18-29

Process Map (Python, Heuristic Miner)

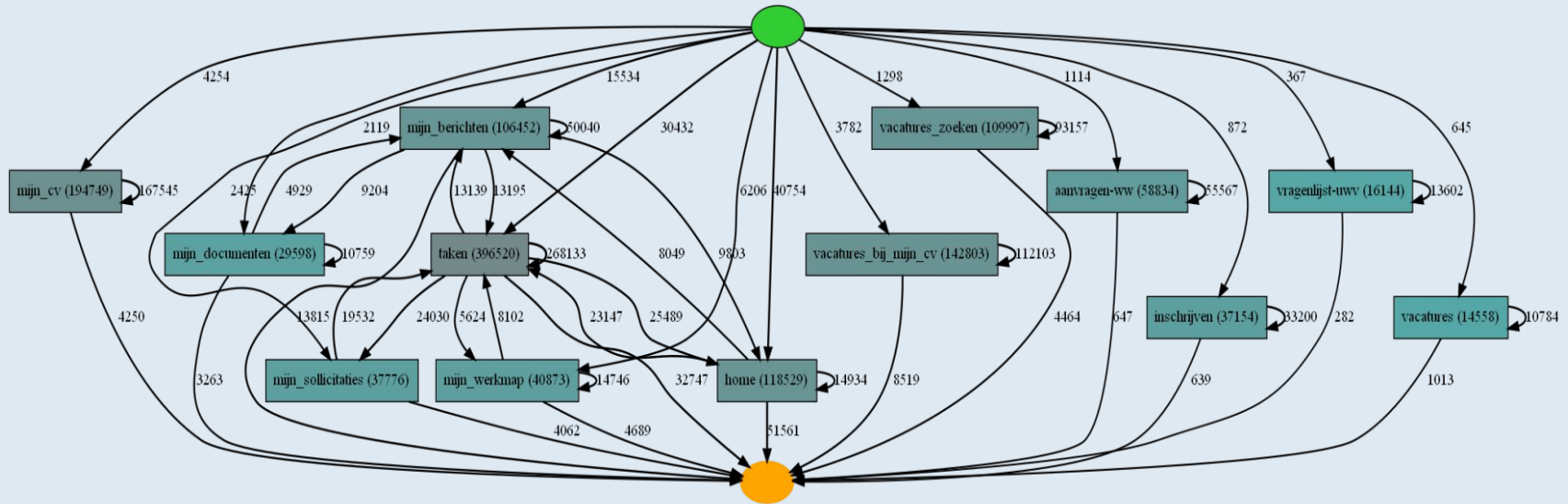


Segment 2: Age 30-39



Segment 2: Age 30-39

Process Map (Python, Heuristic Miner)

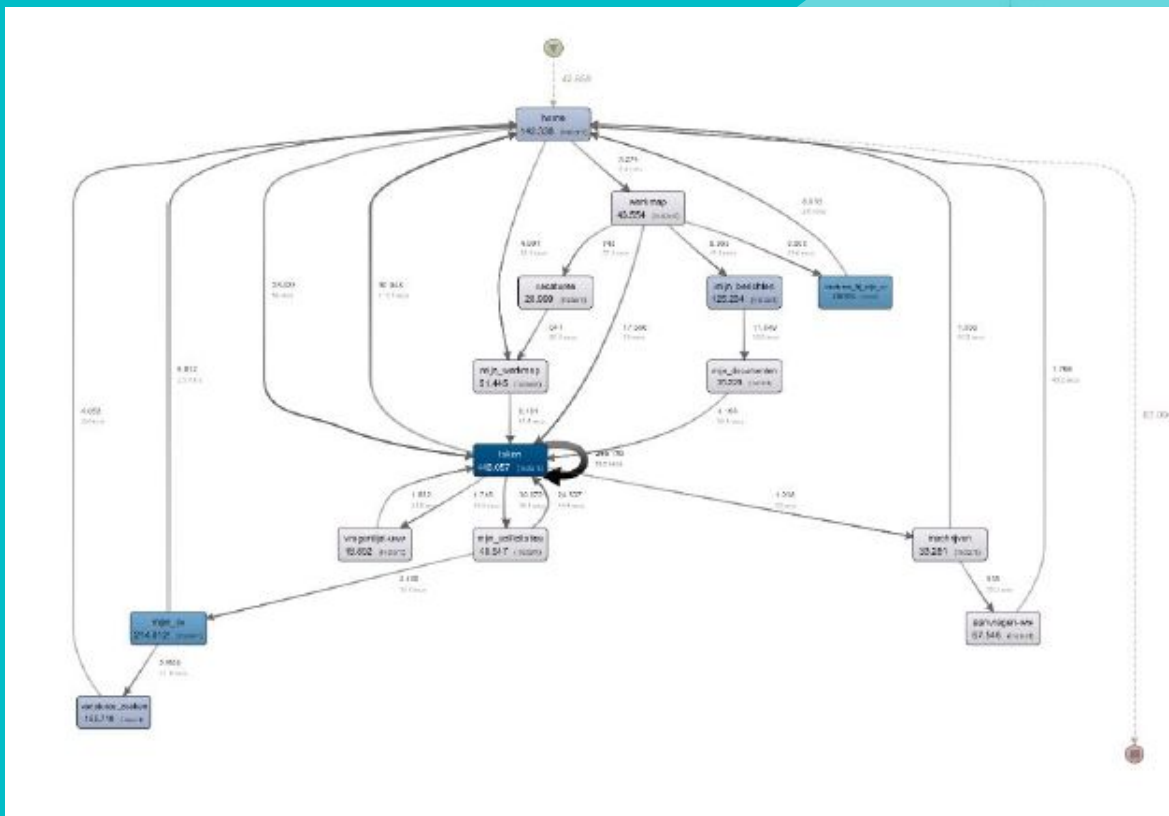


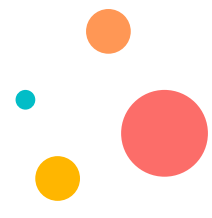


Segment 3: Age 40-49

27

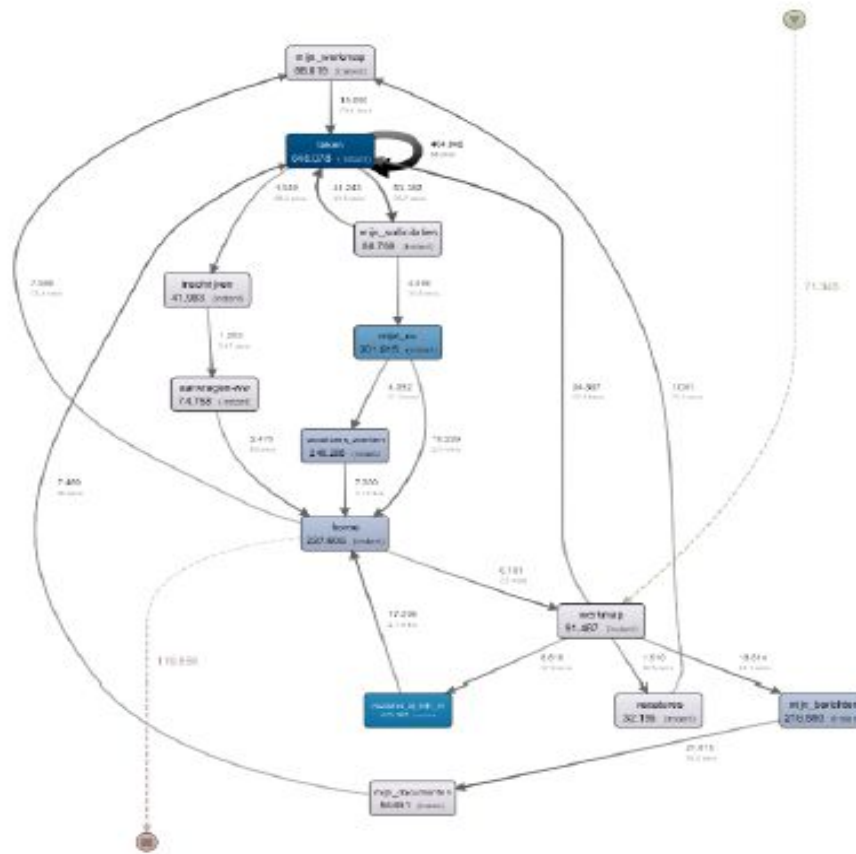
Process Map (Disco)



[illegible]

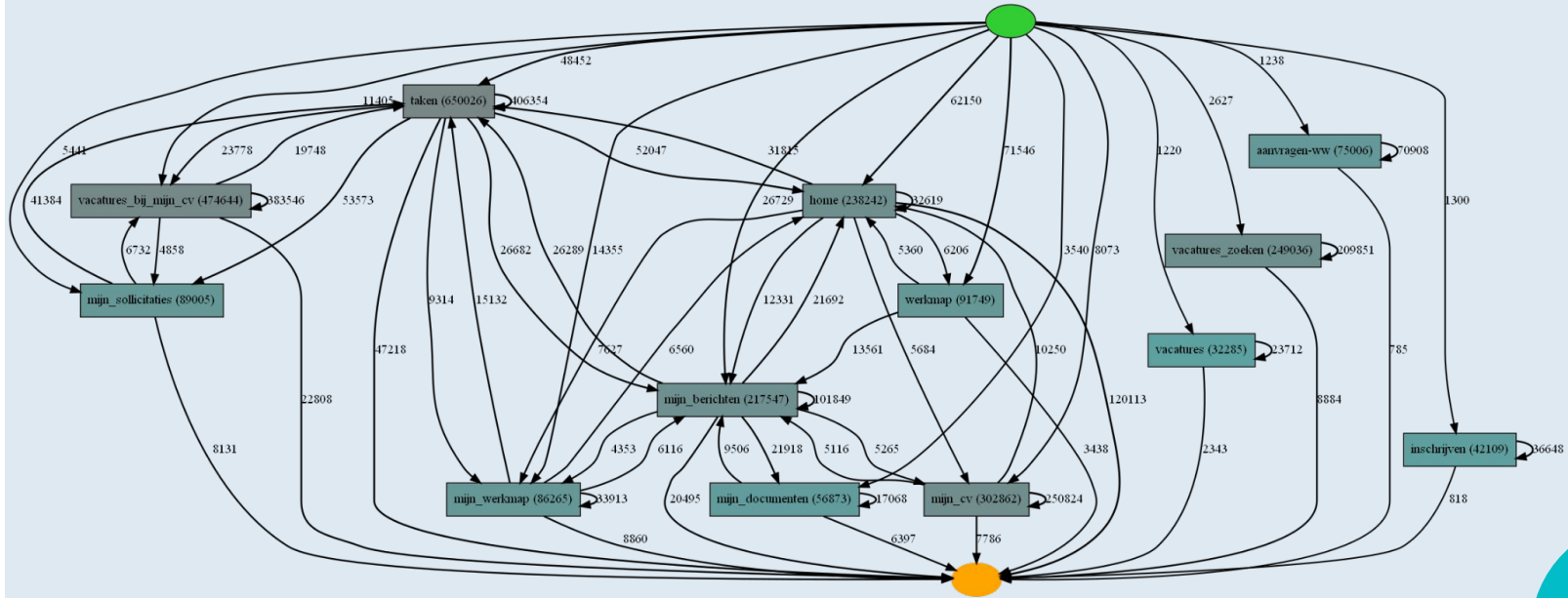
Segment 4: Age 50-65

Process Map (Disco)



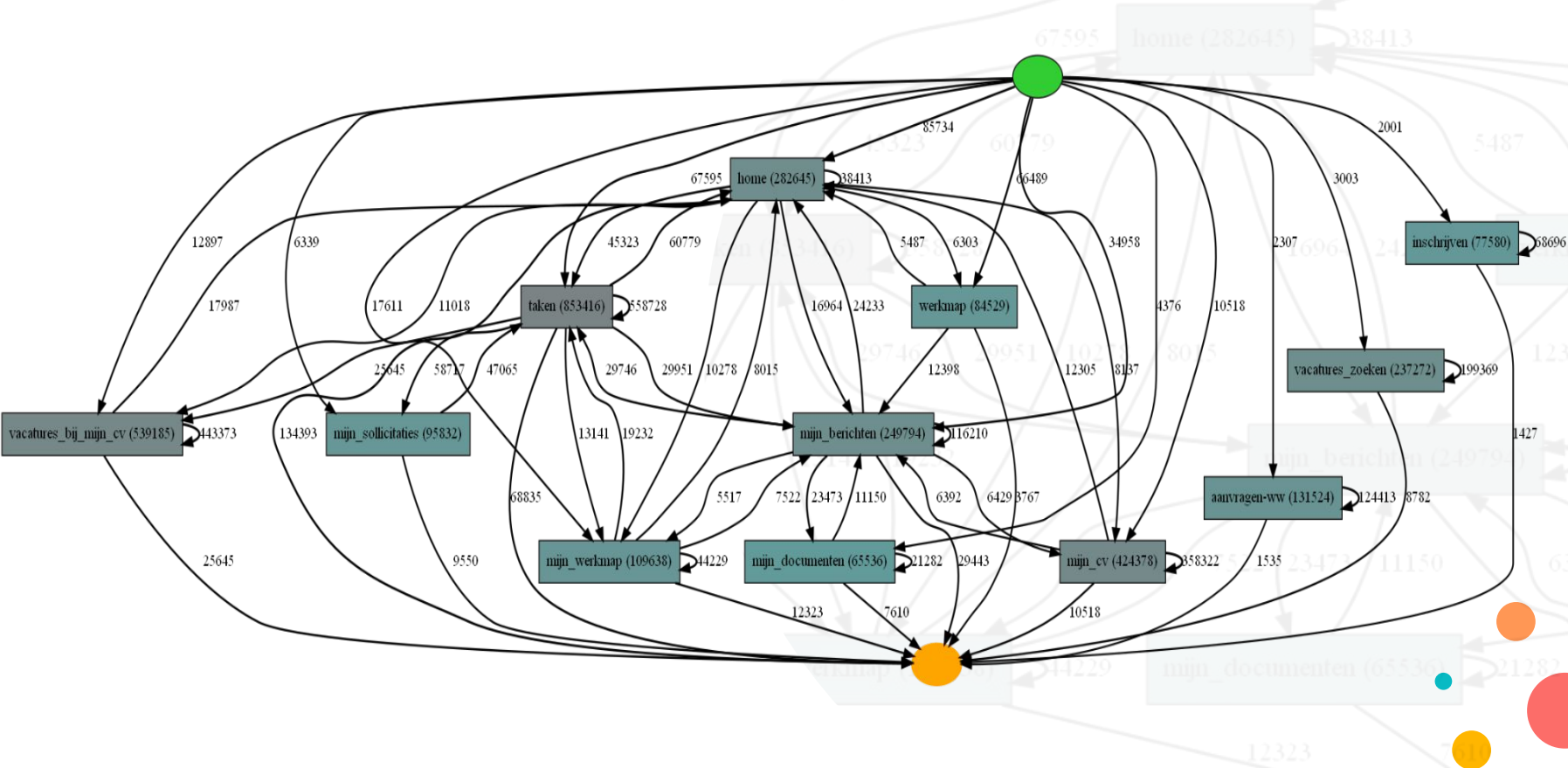
Segment 4: Age 50-65

Process Map (Python, Heuristic Miner)



Segment 6: Male

Process Map (Python, Heuristic Miner)





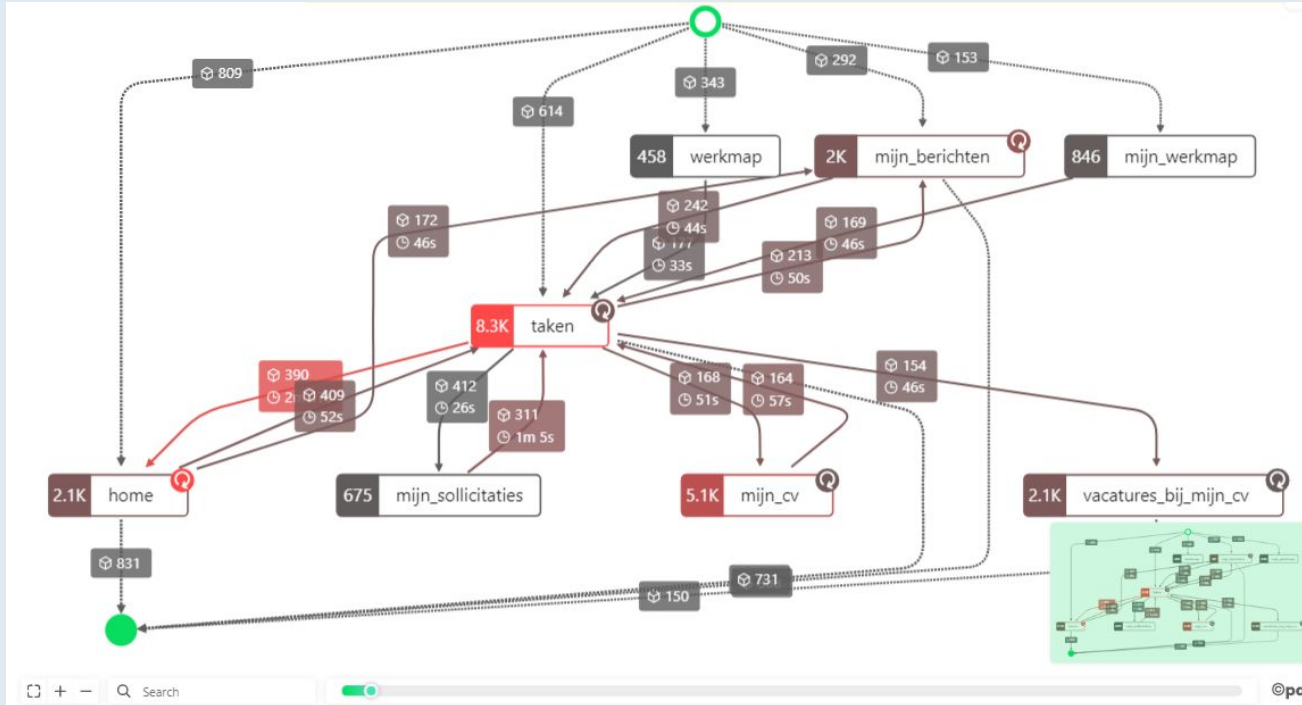
Some Insights:

- While segments 1 to 3 clearly start with activity **home**, this is not true for segment 4. Segment 4 starts with the **werkmap** activity represents the entrance. Assuming that the home page is the intended starting point that a customer is supposed to use when logging into the system, this deviation is remarkable because their primary interest is managing the werkmap tasks, which significantly deviates from other segments.
- The activity vacatures (Dutch for “vacancy”, i. .e an open job offering) is frequent in all segments except for segment 1. Instead, customers from segment 1 do look for vacancies (activity vacatures_zoeken).
- Activit vragenlist-uwv (list of questions) appears in segments 1 to 3 but not in segment 4, which shows that older customers seem to have fewer questions.



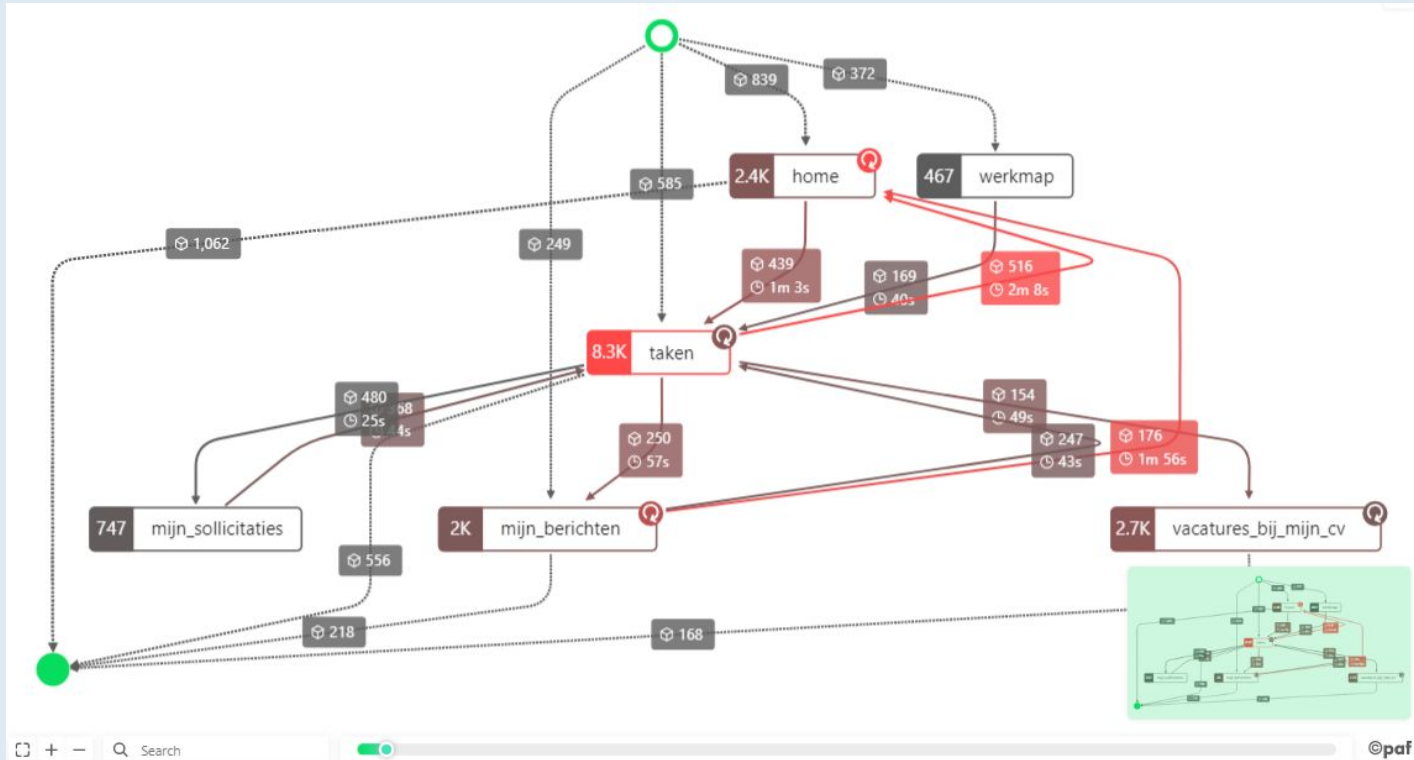
Segment 1: Age 18-29

Process Map (Power BI, Pafnow)



Segment 2: Age 30-39

Process Map (Power BI, Pafnow)



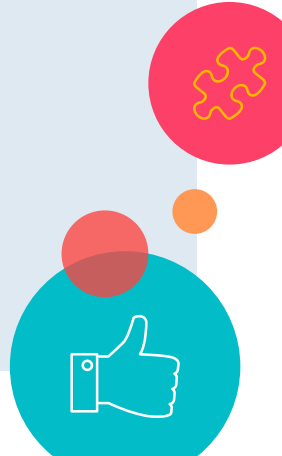


Challenge 2: Changes of Usage Patterns Over Time



Steps:

1. Find out the change pattern in number of **customers** over time
2. Find out the change pattern in number of **clicks** over time
3. Find out the change pattern in **web pages usage** over time
4. Creating process models representing **first** and **fifteenth** sessions with Hursitic Miner



Number of Customers having at least 1, 2, ...and 15 Number of Sessions in website

The following tables introduces the number of customer having at least 15 sessions of using the website. The analysis of results suggests that 26,647 users had at least one session which significantly decreased to 14,800 who had at least 15 sessions. From this statistics, we can infer that the users don't tend to use the website after they used it for the first time and this trends continues over time.

The average drop rate between sessions is about 4 %.

As you can see, the numbers are very similar, and the pattern is the same. There is only a slight difference between some numbers because, as we mentioned before, the authors probably deleted some records in their preprocessing, which we don't know why.

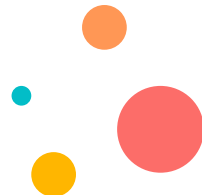
	Session	Customers	Change
0	1 Session	26647	-
1	2 Sessions	25705	-3.54
2	3 Sessions	24733	-3.78
3	4 Sessions	23815	-3.71
4	5 Sessions	22901	-3.84
5	6 Sessions	21959	-4.11
6	7 Sessions	21059	-4.1
7	8 Sessions	20151	-4.31
8	9 Sessions	19298	-4.23
9	10 Sessions	18462	-4.33
10	11 Sessions	17699	-4.13
11	12 Sessions	16970	-4.12
12	13 Sessions	16209	-4.48
13	14 Sessions	15444	-4.72
14	15 Sessions	14800	-4.17

Our Table from python codes

Table 5. Number of customers having at least # of sessions.

Session	Customers	Change
1 Session	26,647	-
2 Sessions	25,705	-3.54%
3 Sessions	24,733	-3.78%
4 Sessions	23,814	-3.72%
5 Sessions	22,896	-3.85%
6 Sessions	21,956	-4.11%
7 Sessions	21,053	-4.11%
8 Sessions	20,143	-4.32%
9 Sessions	19,292	-4.22%
10 Sessions	18,452	-4.35%
11 Sessions	17,691	-4.12%
12 Sessions	16,962	-4.12%
13 Sessions	16,202	-4.48%
14 Sessions	15,435	-4.73%
15 Sessions	14,792	-4.17%

Table from our base Article



Click Trend Over Time

	# of Sessions	# of Click Logs	Change(%)	Average # of Clicks per customer pro session
0	1 Sessions	603405	-	22.64
1	2 Sessions	464535	-23.01	18.07
2	3 Sessions	369503	-20.46	14.94
3	4 Sessions	309310	-16.29	12.99
4	5 Sessions	276435	-10.63	12.07
5	6 Sessions	251105	-9.16	11.44
6	7 Sessions	228086	-9.17	10.83
7	8 Sessions	211034	-7.48	10.47
8	9 Sessions	200614	-4.94	10.40
9	10 Sessions	187336	-6.62	10.15
10	11 Sessions	180325	-3.74	10.19
11	12 Sessions	172422	-4.38	10.16
12	13 Sessions	161150	-6.54	9.94
13	14 Sessions	149654	-7.13	9.69
14	15 Sessions	143238	-4.29	9.68

Our Table from python codes

Table 6. Click trend over time (sessions are sorted).

# of Sessions	# of Click Logs	Change	Average # of Clicks per customer pro session
1. Sessions	453,928	-	17.03
2. Sessions	381,350	-16%	14.84
3. Sessions	343,817	-10%	13.90
4. Sessions	309,035	-10%	12.98
5. Sessions	284,811	-8%	12.44
6. Sessions	266,686	-6%	12.15
7. Sessions	248,288	-7%	11.79
8. Sessions	231,511	-7%	11.49
9. Sessions	213,970	-8%	11.09
10. Sessions	201,291	-6%	10.91
11. Sessions	191,186	-5%	10.81
12. Sessions	183,143	-4%	10.80
13. Sessions	171,892	-6%	10.61
14. Sessions	162,147	-6%	10.51
15. Sessions	151,629	-6%	10.25

Table from our base Article

In the above tables we see the aggregated number of logs per session sorted according to the time. The first column of this table indicates not the amount but the order of the sessions. E.g. the number of click logs in the first session of all individual customers is equal to 603,405. These results also propose that the number of click logs per session drops over time by decreasing to 143,238 in fifteenth session of all customers.

From the last two slides we can argue that, not only the number of customers using the website decreases over time but also the average clicks per session follows a negative trend.

Change of Most Frequent Web Pages Usage Over Time and Sessions

To answer the question of how the usage patterns change, we have aggregated the log data of customers beginning from their first session to the fifteenth session and analyzed the website's visits. The above figure provides valuable insights into changes in website usage behavior over time. We have introduced the visited web pages with a relative frequency higher than 1.1%. From the underlying diagram, we can detect a significant drop in the visit of the "mijn_cv" page. The relative frequency of "**mijn_cv**" decreased to less than 10% in the fifteenth, more than 25% in users' first session. A significant negative trend is also observed in the visit frequency of "**aanvragen-ww**" and "**inschrijven**".

In contrast to **mijn_cv**, "**taken**" page follows an increasing preference trend over time. The relative frequency of taken has increased from 12.32% in the first sessions to 30.14% in the fifteenth sessions. In other web pages such as "**werkmap**", "**mijn_sollicitaties**", "**vacatures_zoeken**", "**mijn_berichten**" and etc. we can observe an increasing trend however, the amplitude of the change is not significantly high.

Chart from our base article

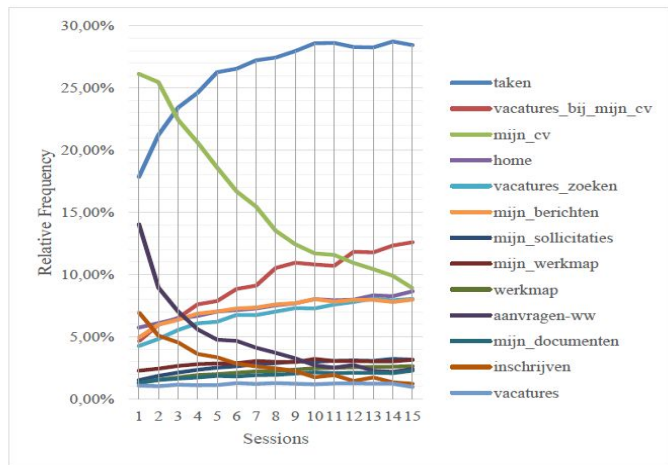
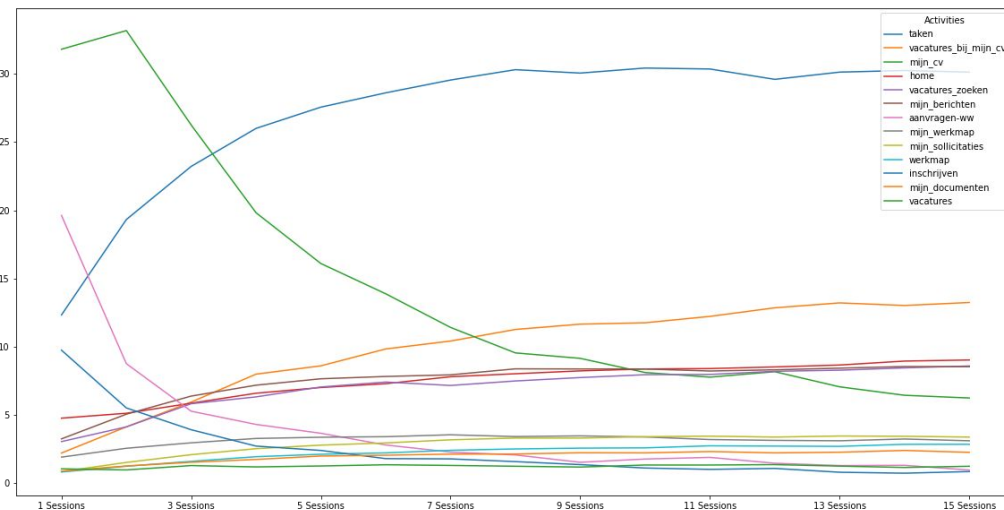


Fig. 5. Change of website usage over time.

Our chart from python codes



Process Models Representing First and Fifteenth Sessions

In order to identify the changes in the transition between websites we need to create the corresponding process maps.

we will create the process map for the 1st and the 15th sessions.

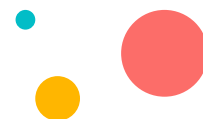
Most Visited Pages In the First and Fifteenth Sessions of Visiting the Website

Most Visited Pages In the First Sessions

	Activity	Absolute Frequency	Relative Frequency(%)
0	mijn_cv	192,033	31.82
1	aanvragen-ww	118,418	19.62
2	taken	74,368	12.32
3	inschrijven	58,738	9.73
4	home	28,596	4.74
5	mijn_berichten	19,438	3.22
6	vacatures_zoeken	18,202	3.02
7	vacatures_bij_mijn_cv	13,170	2.18
8	mijn_werkmap	11,374	1.88
9	vacatures	6,171	1.02
10	mijn_documenten	5,109	0.85
11	ww	4,979	0.83
12	mijn_sollicitaties	4,930	0.82

Most Visited Pages In the Fifteenth Sessions

	Activity	Absolute Frequency	Relative Frequency(%)
0	taken	43,175	30.14
1	vacatures_bij_mijn_cv	18,948	13.23
2	home	12,905	9.01
3	vacatures_zoeken	12,278	8.57
4	mijn_berichten	12,208	8.52
5	mijn_cv	8,916	6.22
6	mijn_sollicitaties	4,794	3.35
7	mijn_werkmap	4,415	3.08
8	werkmap	4,044	2.82
9	mijn_documenten	3,193	2.23
10	vragenlijst-uwv	2,858	2.00
11	vacatures	1,740	1.21



First Sessions Process Map (Disco)

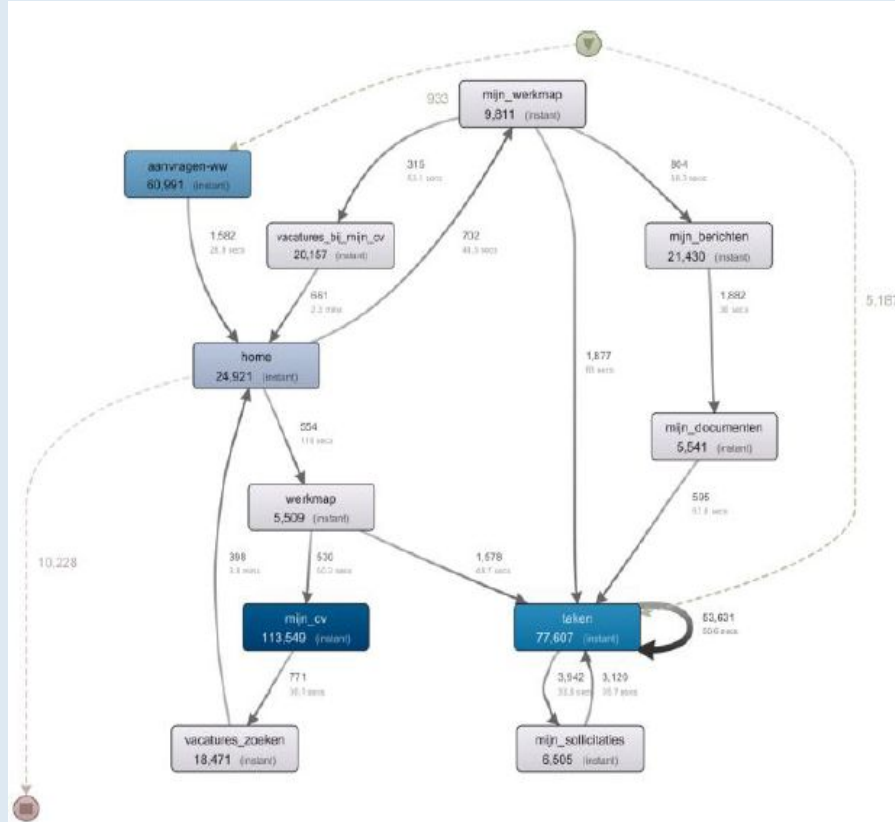
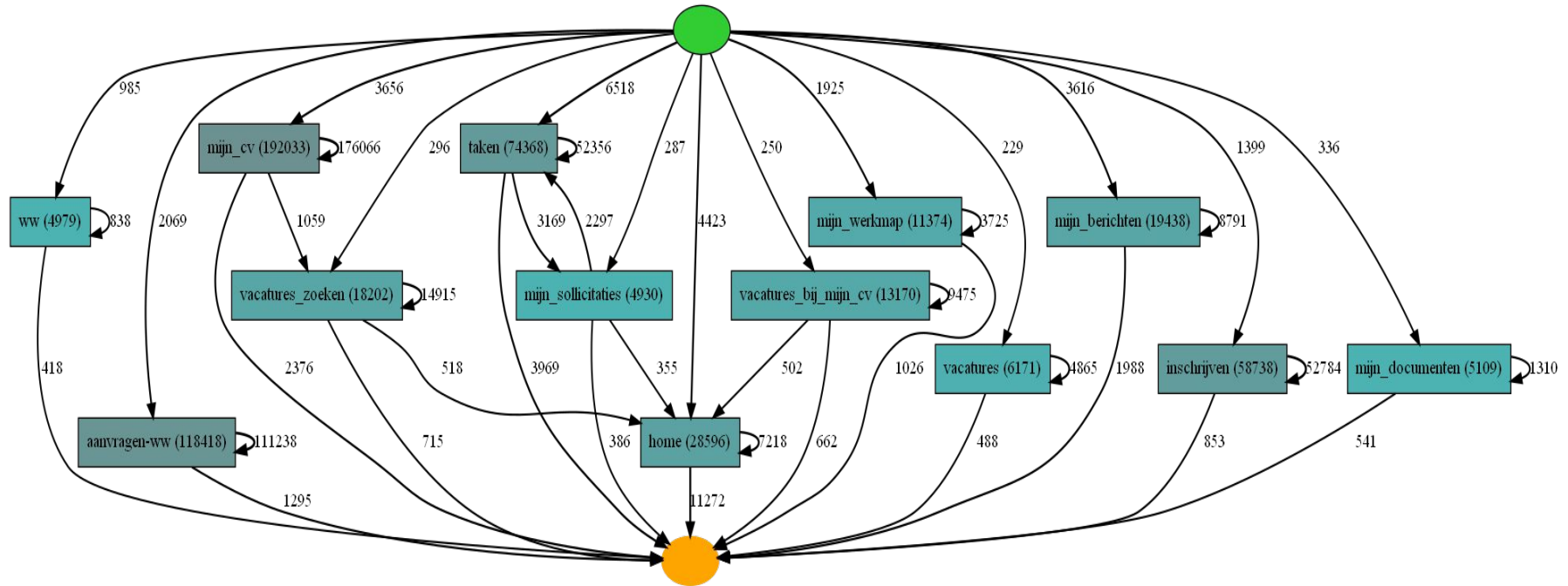


Fig. 6. Process Model representing first sessions of customers.

First Sessions Process Map (Python, Heuristic Miner)



Fifteenth Sessions Process Map (Disco)

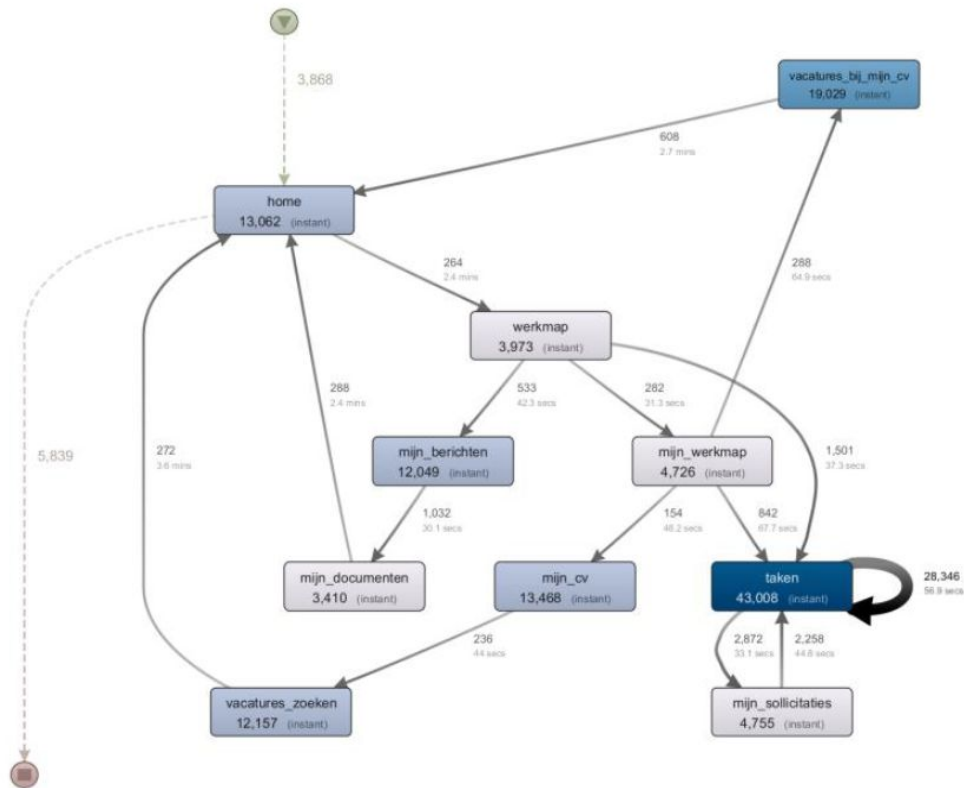
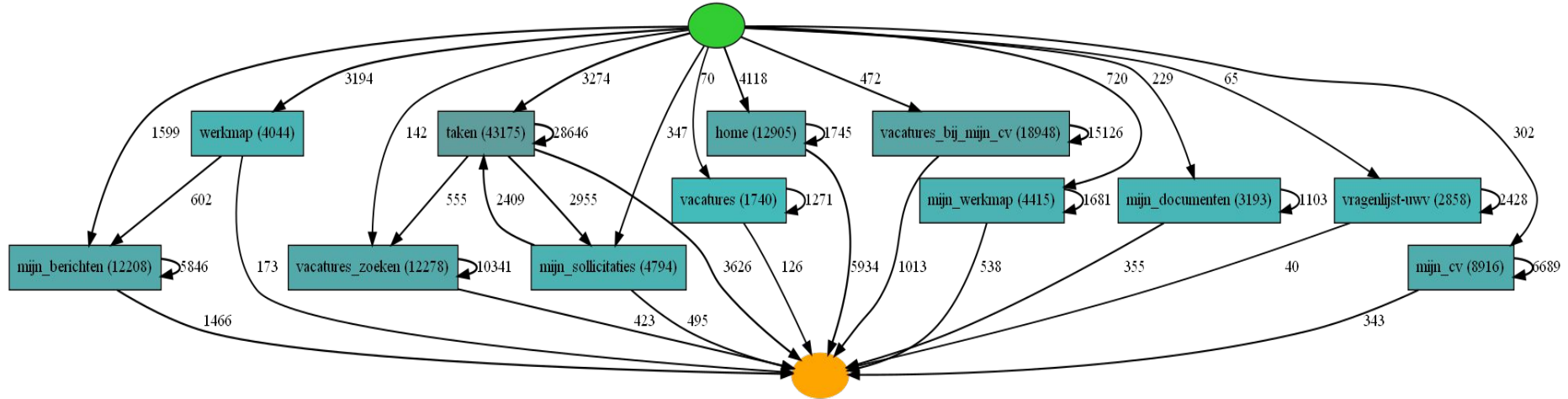
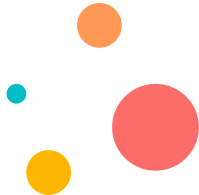


Fig. 7. Process Model representing fifteenth sessions of customers.

Fifteenth Sessions Process Map (Python, Heuristic Miner)



A significant drop in the number of clicks can also be directly observed here compared to the process map of the first sessions we created before. These process diagrams allow us to observe the process paths evolved. A narrow analysis of process models reveals that the sequence of transitions did not change significantly with a few exceptions. **aanvragen-ww** was one of the most visited websites in the earlier sessions, which almost disappeared towards the last sessions (the number of clicks dropped dramatically).





Challenge 4: Communication Channels



Intro:

Within this section, we give answers to some questions: Are customers more likely to use these channels again after they have used them for the first time? How is the process map when we consider the communication channels? After visiting which web pages customers will use the communication channels?

To find the answer the question we need to read the four following datasets analyze them separately and then merge them:

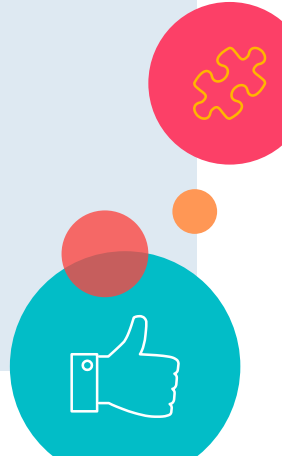
- BPI Challenge 2016: Clicks Logged In
- BPI Challenge 2016: Questions
- BPI Challenge 2016: Werkmap Messages
- BPI Challenge 2016: Complaints





Steps:

1. Usage of communication channels
2. Cleaning and merging datasets
3. Creating Frequency Matrix
4. Creating Performance Matrix
5. Process map of merged data



Total and unique number of communication channel uses, customers who contacted via complaints, werkmap and questions (call center)

Table from python codes

	Communication Channel	Total Use	Unique Customer Use	Relative Use (%)	Averages (Total/Unique)
0	Questions(phone Call)	123,403	21,533	65.03	5.73
1	Werkmap	66,058	16,653	34.81	3.97
2	Complaints	289	226	0.15	1.28
3	Total Communications(sum of the three above)	189,750	-	-	-

Chart from PowerBI analysis

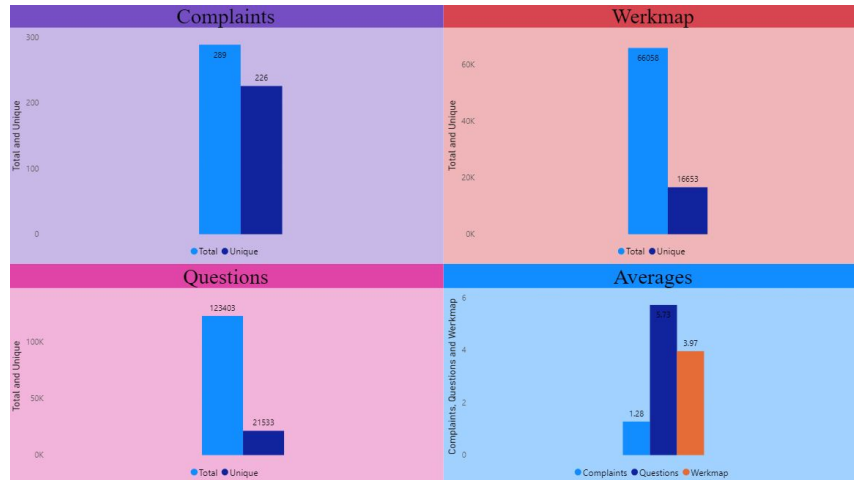


Chart from our base Article

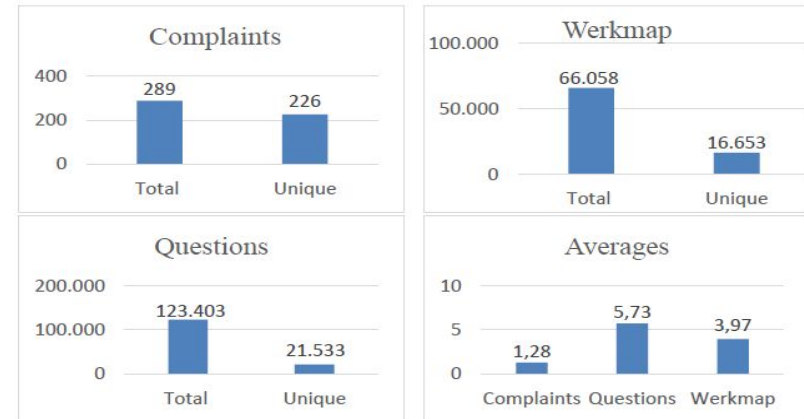


Fig. 9. Total number of communication channel uses. Second columns indicate the number of unique customers who contacted via complaints, werkmap and questions (call center)

Merging Datasets

	customerID	activity	date
9217	1297588	workflow_message	2015-07-01 00:26:11.000
29692	2035791	workflow_message	2015-07-01 00:42:30.000
39687	2016236	workflow_message	2015-07-01 00:43:59.000
59798	2040395	workflow_message	2015-07-01 01:07:41.000
3271608	797444	mijn_berichten	2015-07-01 02:36:59.713
...
121961	602197	phone_call	2016-02-29 17:00:10.000
122567	2106983	phone_call	2016-02-29 17:00:28.000
64831	2237416	phone_call	2016-02-29 17:03:09.000
121246	2106983	phone_call	2016-02-29 17:04:14.000
121673	2106983	phone_call	2016-02-29 17:05:13.000

6758343 rows × 3 columns

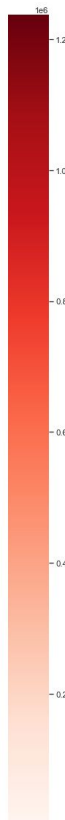
```
## Merging the four simplified datasets
```

```
simple_customer_journey = pd.concat([workflow_messages_simp, phone_calls_simp, complaints_simp, clicks_logged_in_simp])  
simple_customer_journey.sort_values(by='date', ascending=True)
```

Frequency Matrix



workflow_message	2608	2394	8	149	1962	2146	1821	16792	12936	10424	1340	904	5589	5328	219	105
phone_call	453	59419	108	914	1345	2275	3048	9610	13255	12912	1180	515	1138	10505	305	1281
complaint	1	84	28	1	5	6	5	13	41	33	4	2	3	47	1	
inschrijven	46	574	1	131715	153	926	1488	707	5152	2455	221	90	147	680	181	4262
vacatures_bi_mijn_cv	1361	2404	6	279	770180	12592	7343	14816	45493	49859	10037	22713	2249	12219	1166	180
mijn_cv	1246	2051	2	1509	16351	748963	7541	14758	27471	31952	4404	11433	6670	4248	644	749
mijn_werkmap	616	2555	6	959	13600	8366	80923	17074	20279	42948	6293	3211	3925	5136	696	450
mijn_berichten	30336	7312	18	810	13579	13853	12164	247372	55352	70657	10675	5228	48803	9441	1022	468
home	10629	19413	49	3318	26809	21296	30743	56320	153306	137626	13079	7438	8211	75229	8235	2771
taken	6286	13509	36	5942	54601	36598	31514	75325	157549	1237215	128700	24583	12886	29362	3159	868
mijn_solicities	946	1762	8	147	13347	9999	4995	16477	25181	104191	11087	6877	3221	4362	434	112
vacatures_zoeken	593	1155	2	165	21432	5577	3124	4662	20914	20650	4121	493314	854	4848	655	122
mijn_documenten	10444	1826	6	385	4190	6699	3781	24399	16383	19414	4309	1876	45184	3293	188	229
werkmap	210	972	1	319	14910	7765	14455	27258	11838	73733	7644	3664	3766	12002	2919	192
vacatures	102	415	1	102	1270	641	2896	935	6593	3624	500	566	174	2277	58516	66
aanvragen-ww	28	827	1	847	60	325	968	374	8327	946	51	30	49	1057	338	237174
workflow_message	phone_call	complaint	inschrijven	vacatures_bi_mijn_cv	mijn_cv	mijn_werkmap	mijn_berichten	home	taken	mijn_solicities	vacatures_zoeken	mijn_documenten	werkmap	vacatures	aanvragen-ww	



54



Performance Matrix



workflow_message	17521	439975	467	10235	36044	80674	96562	602753	591331	508120	35829	19522	59865	498029	6446	37181
phone_call	36355	984770	15181	257380	136138	589326	447927	1297202	1711892	2181809	152958	67288	149222	1430254	50923	336253
complaint	13	8275	359	9	180	1014	491	2457	6545	8415	3697	18	35	7023	3	
inschrijven	1746	71269	71	15426	4044	26318	17538	36281	59003	95129	6528	2360	8281	39069	1730	15969
vacatures_bi_mijn_cv	4515	252688	411	18534	346575	141701	261074	264091	850059	719614	69122	37986	30528	688757	22487	17803
mijn_cv	3897	297104	120	51282	55739	341473	133188	294767	457704	547242	41119	21081	31841	357042	18026	74206
mijn_werkmap	9700	262957	389	19023	53900	100273	278419	212284	462472	439399	34516	22195	28337	356699	14292	30905
mijn_berichten	24690	762235	3886	17658	88717	165359	190039	975314	938838	1095033	83253	34635	102724	792707	14763	51002
home	50208	2795365	7062	164038	606364	935468	1152107	2483233	7217183	5417295	458965	169878	290560	6277775	107226	196514
taken	28475	1435400	5028	65605	291429	329027	571883	1262728	3215696	4640651	363329	76627	146514	2582908	46029	56524
mijn_solicities	3408	196271	463	13100	46430	36601	85027	181217	435762	556581	76347	13960	22603	363109	7350	16300
vacatures_zoeken	1675	157978	79	24500	74231	84330	94260	134638	418073	371920	31163	91748	17165	311724	15937	24342
mijn_documenten	7118	186761	93	12342	23278	52202	60515	183568	244970	296875	23543	8550	48145	250597	5089	24005
werkmap	4517	92600	6	18916	15642	26120	30224	73838	95299	139593	10022	5446	9666	204136	2381	14551
vacatures	1103	68482	13	6101	11785	36581	37427	61678	115035	129543	12039	9724	8882	97790	11822	9250
aanvragen_vw	82	95714	62	13764	3819	30307	25817	68804	64264	135964	6373	2396	8770	43839	2174	26550
workflow_message																
phone_call																
complaint																
inschrijven																
vacatures_bi_mijn_cv																
mijn_cv																
mijn_werkmap																
mijn_berichten																
home																
taken																
mijn_solicities																
vacatures_zoeken																
mijn_documenten																
werkmap																
vacatures																
aanvragen_vw																

1e6

7

6

5

4

3

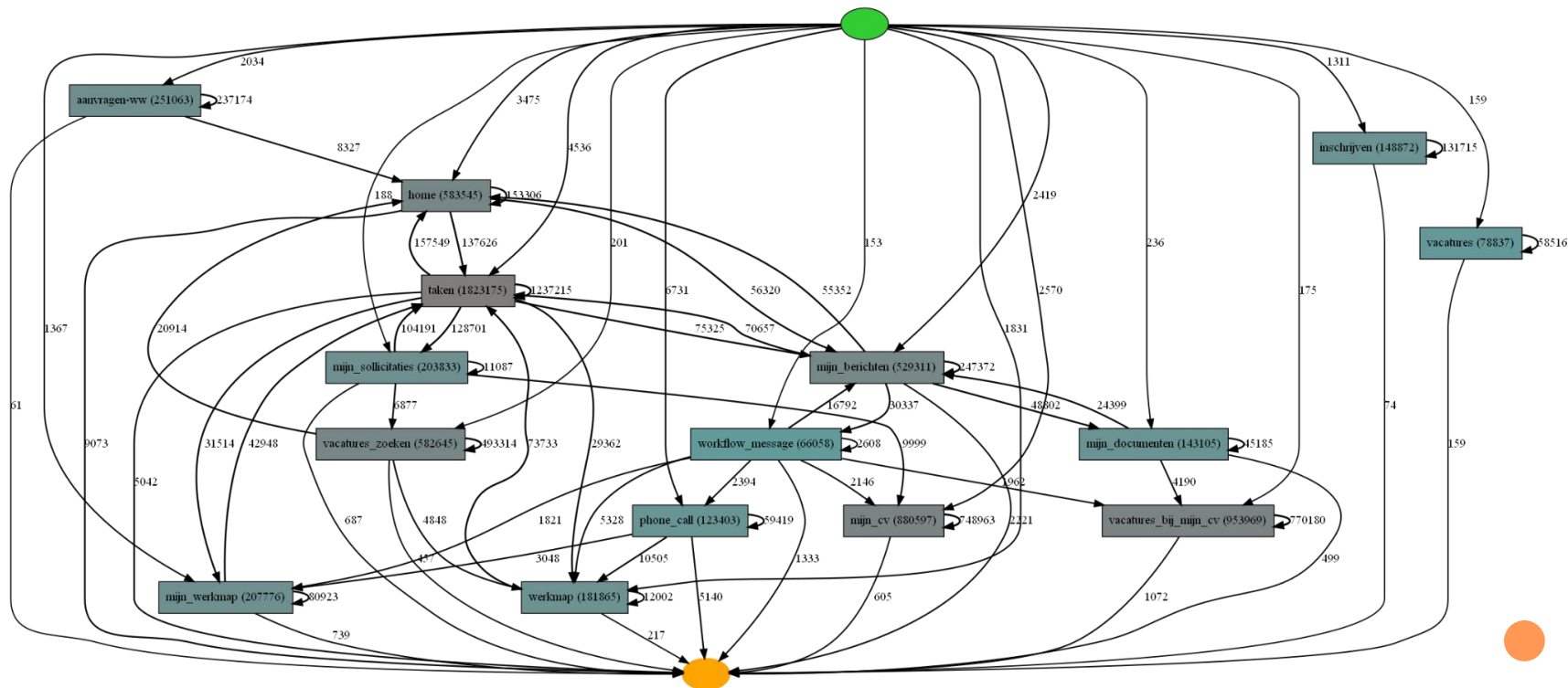
2

1

55



Merged data Process Map (Python, Heuristic Miner)





Thanks!

Any questions?

You can find us at:

- sinaaghaee21@gmail.com
 - farzanehmahmoudi870@gmail.com
 - varchandi.sahar@gmail.com
- 