

Checkout Durations for Books in Different Dewey Classes

Zhuowei Cheng

This assignment focuses on the duration of checkout time for books in different dewey classes and tries to answer the following questions:

1. How does the duration of checkout times differ for books in different dewey classes over the past 15 years.
2. Does the covid situation impact the checkout durations?
3. For the dewey classes with the longest and shortest checkout durations, which subclasses contribute the most to the results.

MySQL query:

There are two queries used in this assignment with some parameter changes. One query is to get the average checkout durations for books in different dewey classes and the other one is to get the number of checkouts for books in different dewey classes. Here's the example query:

Query 1:

This example query outputs the average checkout durations for books in different dewey classes (from 000 to 900) from 2006-2020. I also looked at the subclasses in dewey 400,700 and 800. (To look at the subclasses, change the value in blue to the target classes.)

```
SELECT
  SUM(CASE
    WHEN deweyClass > 000 AND deweyClass < 100 THEN TIMESTAMPDIFF(HOUR, cout, cin) / 24.0
    ELSE 0
  END) / SUM(CASE
    WHEN deweyClass > 000 AND deweyClass < 100 THEN 1
    ELSE 0
  END) AS 'dewey_000',
  SUM(CASE
    WHEN deweyClass > 100 AND deweyClass < 200 THEN TIMESTAMPDIFF(HOUR, cout, cin) / 24.0
    ELSE 0
  END) / SUM(CASE
    WHEN deweyClass > 100 AND deweyClass < 200 THEN 1
    ELSE 0
  END) AS 'dewey_100',
  SUM(CASE
    WHEN deweyClass > 200 AND deweyClass < 300 THEN TIMESTAMPDIFF(HOUR, cout, cin) / 24.0
    ELSE 0
  END) / SUM(CASE
```

```

    WHEN deweyClass > 200 AND deweyClass < 300 THEN 1
    ELSE 0
END) AS 'dewey_200',
SUM(CASE
    WHEN deweyClass > 300 AND deweyClass < 400 THEN TIMESTAMPDIFF(HOUR, cout, cin) / 24.0
    ELSE 0
END) / SUM(CASE
    WHEN deweyClass > 300 AND deweyClass < 400 THEN 1
    ELSE 0
END) AS 'dewey_300',
SUM(CASE
    WHEN deweyClass > 400 AND deweyClass < 500 THEN TIMESTAMPDIFF(HOUR, cout, cin) / 24.0
    ELSE 0
END) / SUM(CASE
    WHEN deweyClass > 400 AND deweyClass < 500 THEN 1
    ELSE 0
END) AS 'dewey_400',
SUM(CASE
    WHEN deweyClass > 500 AND deweyClass < 600 THEN TIMESTAMPDIFF(HOUR, cout, cin) / 24.0
    ELSE 0
END) / SUM(CASE
    WHEN deweyClass > 500 AND deweyClass < 600 THEN 1
    ELSE 0
END) AS 'dewey_500',
SUM(CASE
    WHEN deweyClass > 600 AND deweyClass < 700 THEN TIMESTAMPDIFF(HOUR, cout, cin) / 24.0
    ELSE 0
END) / SUM(CASE
    WHEN deweyClass > 600 AND deweyClass < 700 THEN 1
    ELSE 0
END) AS 'dewey_600',
SUM(CASE
    WHEN deweyClass > 700 AND deweyClass < 800 THEN TIMESTAMPDIFF(HOUR, cout, cin) / 24.0
    ELSE 0
END) / SUM(CASE
    WHEN deweyClass > 700 AND deweyClass < 800 THEN 1
    ELSE 0
END) AS 'dewey_700',
SUM(CASE
    WHEN deweyClass > 800 AND deweyClass < 900 THEN TIMESTAMPDIFF(HOUR, cout, cin) / 24.0
    ELSE 0
END) / SUM(CASE
    WHEN deweyClass > 800 AND deweyClass < 900 THEN 1
    ELSE 0

```

```

END) AS 'dewey_800',
SUM(CASE
  WHEN deweyClass > 900 AND deweyClass < 1000 THEN TIMESTAMPDIFF(HOUR, cout, cin) / 24.0
  ELSE 0
END) / SUM(CASE
  WHEN deweyClass > 900 AND deweyClass < 1000 THEN 1
  ELSE 0
END) AS 'dewey_900'
FROM
  spl_2016.inraw
WHERE
  itemtype LIKE '%bk' AND YEAR(cout) >= '2006' YEAR(cout) <= '2020'
GROUP BY YEAR(cout);

```

Query 2:

This example query outputs the number of checkouts for different dewey subclasses (800-890) from 2006-2020. I also looked at the subclasses in dewey 400 and 700. (To look at 400 and 700, I changed the value in blue to the subclasses I wanted to look at.)

```

SELECT
  SUM(CASE
    WHEN deweyClass > 800 AND deweyClass < 810 THEN 1
    ELSE 0
  END) AS 'dewey_000',
  SUM(CASE
    WHEN deweyClass > 810 AND deweyClass < 820 THEN 1
    ELSE 0
  END) AS 'dewey_100',
  SUM(CASE
    WHEN deweyClass > 820 AND deweyClass < 830 THEN 1
    ELSE 0
  END) AS 'dewey_200',
  SUM(CASE
    WHEN deweyClass > 830 AND deweyClass < 840 THEN 1
    ELSE 0
  END) AS 'dewey_300',
  SUM(CASE
    WHEN deweyClass > 840 AND deweyClass < 850 THEN 1
    ELSE 0
  END) AS 'dewey_400',
  SUM(CASE
    WHEN deweyClass > 850 AND deweyClass < 860 THEN 1
    ELSE 0
  END) AS 'dewey_500',
  SUM(CASE

```

```

    WHEN deweyClass > 860 AND deweyClass < 870 THEN 1
    ELSE 0
  END) AS 'dewey_600',
  SUM(CASE
    WHEN deweyClass > 870 AND deweyClass < 880 THEN 1
    ELSE 0
  END) AS 'dewey_700',
  SUM(CASE
    WHEN deweyClass > 880 AND deweyClass < 890 THEN 1
    ELSE 0
  END) AS 'dewey_800',
  SUM(CASE
    WHEN deweyClass > 890 AND deweyClass < 900 THEN 1
    ELSE 0
  END) AS 'dewey_900'
FROM
  spl_2016.inraw
WHERE
  itemtype LIKE '%bk' AND YEAR(cout) >= '2006' YEAR(cout) <= '2020'
GROUP BY YEAR(cout);

```

Output from the query:

1. Checkout durations (days) for different dewey classes.

	dewey_000	dewey_100	dewey_200	dewey_300	dewey_400	dewey_500	dewey_600	dewey_700	dewey_800	dewey_900
2006	49.5014493	53.1280684	65.978756	61.7768753	62.873533	63.3208433	53.6484484	46.0813565	71.5579067	63.2900575
2007	48.7105879	50.007625	61.7578949	59.0962262	60.6269313	60.5373129	52.108761	43.8848794	65.5400707	60.069222
2008	43.9878569	44.4118134	53.9774414	50.5442733	53.162138	53.4880928	44.9829247	37.9058771	56.108403	53.4251366
2009	42.4142349	43.7689063	52.7048532	48.1683027	51.1637771	49.3002865	42.5949448	35.9064826	52.3652499	50.5857075
2010	41.5490873	41.5395504	46.4350559	42.4736343	44.855347	43.8474708	40.748155	34.2371041	44.10638	44.0253039
2011	30.9859704	30.6504506	30.9883294	29.5581535	32.3368284	29.3022046	30.2936742	24.1624564	28.4248273	28.8605179
2012	27.8540936	25.3002295	26.0611412	25.116698	28.4925837	24.7969775	25.8203987	19.971203	24.1851103	24.0978979
2013	27.2451663	24.7208373	24.7384073	24.6171559	28.3482644	25.0615209	24.754376	19.3048144	23.9445597	23.5026638
2014	26.6894679	24.649897	25.0864891	25.0940506	28.6982453	23.9362527	24.9442995	19.6457476	24.8298932	24.0567962
2015	26.2759169	24.3152441	24.4261977	24.9612985	28.074019	23.5038966	24.6100995	19.4650488	24.2029307	23.6037359
2016	25.5219675	24.4017764	24.8185519	24.4001041	27.7851395	23.2704942	24.4732267	19.6470168	24.5463616	23.6203605
2017	24.5379477	23.8429481	23.8964522	23.5495269	26.729548	22.6662569	23.8968331	18.8632425	23.7357608	22.607804
2018	27.2569849	25.8844212	26.369932	25.8823361	30.1514562	24.5204266	25.7366993	20.4787867	25.6866657	24.3136148
2019	27.1271193	26.3017825	24.9568077	25.0479984	29.4889996	24.3619637	25.506038	21.9313394	24.896897	23.766683
2020	95.8946112	83.6994872	81.0518867	79.0608889	99.5920686	89.0057826	84.5333253	75.7860252	75.8934738	77.5543589

2. Checkout durations(days) for subclasses in class 400 (Language)

	dewey_400	dewey_410	dewey_420	dewey_430	dewey_440	dewey_450	dewey_460	dewey_470	dewey_480	dewey_490
2006	78.9275535	67.3163712	59.9744464	64.6740236	48.4709932	52.2166826	44.5359947	84.5036541	83.1210122	76.1229168
2007	59.6857348	60.4330273	57.682547	60.3643569	45.7006939	46.2847998	51.8045527	75.3901687	99.546696	74.9285941
2008	45.3593192	50.9325677	51.3360816	67.2976195	49.6417144	52.9611448	46.1486583	46.6842377	66.7844068	64.062608
2009	58.7649092	52.7090433	49.3885559	62.4908109	45.2686217	52.5693016	42.2458853	67.4484042	74.1674389	57.0751032
2010	52.0618794	48.4386115	40.9096044	49.3841621	42.9155434	43.0296755	43.0467879	63.247988	67.9485844	53.6065848
2011	35.314488	32.5819704	30.5128242	35.7373638	33.4523271	34.7124074	34.3712529	35.2889597	34.1069132	34.6222673
2012	27.9507066	27.9823968	27.5497471	30.2457327	30.2569932	33.1622575	29.065501	29.2516525	30.1933333	29.6529505
2013	27.9633587	30.0366706	27.8605863	29.3889832	28.784397	29.8552236	27.5220045	28.6283482	24.845024	29.5703797
2014	24.7950184	27.5807191	29.2651239	28.8634526	29.5253597	31.3604283	27.6873305	25.827499	23.2882275	28.4774143
2015	29.2131547	25.8322744	27.732754	31.0294229	28.2676945	29.2360666	29.2762025	32.4505903	28.4367506	28.6701792
2016	26.4533078	25.1699153	27.7251223	29.7934898	28.4240231	30.1645514	28.4802689	24.2350519	22.9396359	28.6594449
2017	22.5741075	23.7037618	26.9916931	29.7903134	27.3473114	28.4227857	26.6503278	30.5496083	17.0704405	27.2976268
2018	27.5371588	31.0957967	30.4548014	32.7958646	29.5424628	30.079738	28.9669956	26.9105854	29.3050245	30.0774851
2019	26.8062218	27.5471924	29.5703831	30.2333328	29.1393528	30.2362285	30.8725775	26.7240475	24.0261581	29.6587931
2020	90.946941	98.5953635	98.2071465	113.808697	111.054403	97.3353219	106.138392	110.953435	69.0654786	94.9499568

3. The number of checkouts for subclasses in class 400 (Language)

	dewey_400	dewey_410	dewey_420	dewey_430	dewey_440	dewey_450	dewey_460	dewey_470	dewey_480	dewey_490
2006	643	1754	9808	991	1323	569	1991	194	115	4420
2007	933	2600	10303	787	1320	437	2121	217	149	4370
2008	1597	2657	13370	840	1219	489	1999	377	132	4453
2009	1020	2700	14280	857	1631	683	2745	214	162	4759
2010	839	2322	14132	964	1966	615	3181	166	141	4869
2011	524	1866	11586	719	1686	726	2480	154	129	3584
2012	366	1077	9173	508	1138	463	1797	101	75	2725
2013	315	1251	9040	435	1049	453	1725	112	196	2480
2014	261	1036	7551	350	953	406	1603	100	218	2201
2015	190	944	7200	354	909	326	1350	113	83	1850
2016	257	962	6460	343	793	315	1210	131	78	1800
2017	280	1030	5889	277	763	286	1145	84	42	1643
2018	194	752	4412	288	678	313	1169	48	53	1439
2019	312	935	5898	460	868	354	1549	61	43	2101
2020	139	277	1511	115	229	105	476	17	14	572

4. Checkout durations (days) for subclasses in class 700 (Arts & recreation)

	dewey_700	dewey_710	dewey_720	dewey_730	dewey_740	dewey_750	dewey_760	dewey_770	dewey_780	dewey_790
2006	77.1498791	50.6240507	55.0155673	75.9379763	31.3635472	68.9784894	81.0326878	56.5711978	73.9314468	70.5000579
2007	73.7312866	49.6896771	57.7848161	70.3854697	31.2529639	69.465075	80.3964878	52.3712218	71.3970044	67.4748014
2008	65.0877808	48.2055714	50.734819	62.5339555	28.8859142	58.0228386	74.3921839	46.7456958	57.6370517	58.7070956
2009	57.2248786	43.9170146	50.0152067	55.0051645	28.2867916	55.3300016	58.7077531	43.3694224	55.268111	53.2451939
2010	48.087381	41.6815504	43.361198	47.1348439	28.6961668	47.4349863	48.8455988	41.26894	47.519711	46.0506288
2011	32.5825802	28.8564454	29.4965953	31.5894667	21.6628504	31.5434919	30.3205335	30.0105446	29.3595556	29.4822039
2012	27.8066911	25.4276775	24.3855456	24.7815188	17.8113683	25.9900407	26.6709187	25.6980032	25.4494471	24.7619816
2013	24.9799562	24.4441271	23.2318099	24.9579673	17.254234	25.3334782	25.1630806	25.172681	24.67082	24.3798499
2014	26.9334331	24.682607	23.0039196	25.1529882	17.5393068	26.11469	24.3968787	25.3297722	25.6022314	24.4553974
2015	26.6168869	25.1685286	23.1831821	24.0964905	17.486616	25.292395	23.718036	24.5167216	25.2743496	24.0380915
2016	25.6960639	26.3957276	22.5267484	24.8736861	17.9072388	25.006601	26.5138095	24.1746414	24.3111647	23.8715764
2017	25.0119333	24.6022057	22.5425378	23.7428635	17.323274	23.8203756	21.9908473	24.433709	23.3233145	22.4827917
2018	27.0557275	25.5009279	24.0989547	27.0439969	18.7274694	26.9139302	26.169795	26.1332547	26.9843129	24.5275688
2019	26.1619187	25.1304091	24.3604636	24.8456934	19.9613129	25.8826915	22.3691075	25.4783027	25.5507712	24.1673044
2020	82.7613566	77.67916	79.1726302	90.9227017	70.5577845	84.4067078	67.1043296	82.3834589	78.678805	83.303502

5. The number of checkouts for subclasses in class 700 (Arts & recreation)

	dewey_700	dewey_710	dewey_720	dewey_730	dewey_740	dewey_750	dewey_760	dewey_770	dewey_780	dewey_790
2006	14783	10280	18147	9520	281159	16128	2580	13452	28923	71437
2007	15697	10611	19592	9527	353629	16872	2606	16478	32349	76523
2008	18239	11286	19877	11056	494536	19291	2844	20854	40273	87550
2009	18479	11025	19654	12457	540678	20026	3162	22914	40665	96573
2010	18426	9781	18198	12956	496960	18759	3363	21161	38495	94629
2011	12784	7559	13341	9809	429504	15061	2093	14604	32920	74369
2012	9182	4363	9146	6171	306704	10309	1421	8898	21350	52748
2013	10249	4165	9410	6275	340093	11192	1325	9962	22658	55148
2014	10313	3717	8981	5800	316562	9942	1156	8788	20423	53940
2015	10164	3669	8492	4921	314350	10070	1151	8051	19307	53943
2016	9397	2975	7393	4730	299376	8742	1065	7189	18871	50796
2017	7800	2358	6121	3632	272292	7634	897	5672	16536	45595
2018	6448	2020	4745	3094	225349	6124	772	4553	11889	37741
2019	9067	2426	6312	4500	155996	8586	891	5936	16002	51841
2020	3247	987	2487	1422	49234	3071	379	2187	5959	18943

6. Checkout durations (days) for subclasses in class 800 (Literature)

	dewey_800	dewey_810	dewey_820	dewey_830	dewey_840	dewey_850	dewey_860	dewey_870	dewey_880	dewey_890
2006	70.2431951	79.0270135	106.68021	112.203439	107.341009	108.517806	81.5787169	120.056137	120.23542	39.320739
2007	64.4945139	72.0676378	89.1127471	97.2952822	90.1357749	86.1848074	83.5022415	79.3204078	97.9272193	38.4536878
2008	58.021599	58.5828767	72.8337811	88.9665116	77.9439631	84.7392001	62.4371245	80.6925441	79.2147264	35.9115023
2009	54.994858	53.7637608	67.201642	69.8072056	68.690386	74.3655349	59.0499462	80.855469	68.6941574	34.3188367
2010	48.7646354	44.6473446	53.9157516	50.6930968	54.316032	54.3627319	48.9154318	65.621902	52.6748634	31.2287988
2011	33.552314	28.7873976	30.5858214	35.1634281	27.7129484	30.6208309	31.2004069	29.6534634	32.8031704	22.6055266
2012	28.6043437	23.8723821	26.1505022	26.9730524	25.6499892	31.3842072	25.4054068	24.6011511	30.9775099	21.1512668
2013	27.011802	23.8537043	25.1175033	26.498838	26.581879	26.2978493	26.5044892	26.1038706	29.1678089	20.8180804
2014	28.5499775	24.0998921	25.7455241	27.8964482	27.9824978	24.1247193	28.0337599	33.9358687	26.8718535	22.3279439
2015	27.5842539	23.7915318	24.9878306	25.2709286	25.3006134	23.720043	27.5922223	26.7431406	27.3143603	21.5408489
2016	28.6567551	24.318394	23.2517679	26.8805831	24.7218512	26.2964777	25.4235425	25.8039122	29.2859469	22.3670247
2017	27.0713218	23.6799982	24.3514095	25.2836921	23.8429002	26.2649704	24.9842825	25.9802936	24.0769589	20.8696755
2018	29.6504307	25.7765133	23.749287	24.1882813	26.1764728	20.9951734	27.5877215	24.1918961	24.261299	23.7756866
2019	28.9871365	23.9282451	24.2564387	24.4950332	27.9181933	29.7680068	28.905623	28.6168669	26.9668401	24.0042032
2020	90.0076728	73.6887345	75.220426	66.7679974	85.3267351	94.8591549	79.5578755	67.7349738	87.4268929	71.201194

7. The number of checkouts for subclasses in class 800 (Literature)

	dewey_800	dewey_810	dewey_820	dewey_830	dewey_840	dewey_850	dewey_860	dewey_870	dewey_880	dewey_890
2006	17776	52582	14114	1847	2318	667	4785	432	1266	36854
2007	19969	57876	14721	2005	2583	820	4666	461	1216	36323
2008	23650	72858	18041	2038	2708	841	5894	673	1799	38210
2009	23460	74621	19074	2306	3046	797	6547	512	1884	35553
2010	21673	72464	17542	2226	3297	900	5872	538	1896	33311
2011	15496	55525	12256	1467	2852	680	4400	363	1072	26334
2012	9628	39591	9365	1104	1574	362	3079	311	717	21631
2013	10715	42034	10183	1112	1520	465	2905	282	877	19657
2014	10638	39709	9785	1051	1488	445	3020	217	861	16205
2015	10229	37645	8461	1292	1276	370	2667	249	784	15070
2016	9812	34571	9919	985	1239	381	2539	262	714	13155
2017	8034	32088	7291	857	1115	398	2211	203	608	12268
2018	6552	26692	6488	879	939	259	2108	180	671	8570
2019	8348	37655	7817	1116	1253	354	2592	251	774	11600
2020	2699	14036	2963	382	442	142	1010	61	326	3316

Data Analysis:

Fig 1 shows the checkout durations for different dewey classes in the past 15 years. Year-wise, from 2006 to 2012, the duration of checkout times gradually reduced. From 2012-2019, the checkout durations have been mostly stable at a range of 20-30 days. Year 2020 has the longest duration across all years because of the covid situation. In 2020, seattle public library posted on its website that “We are temporarily extending the time before we suspend your account for overdue items to 31 days.” and “all returned items will be quarantined – and will remain on your account – for a minimum of 3 days before being checked in.” This can be the reason why we are seeing the big checkout duration up jump. There might also be a policy change from 2006 to 2012. However, I was not able to find relevant information on that.

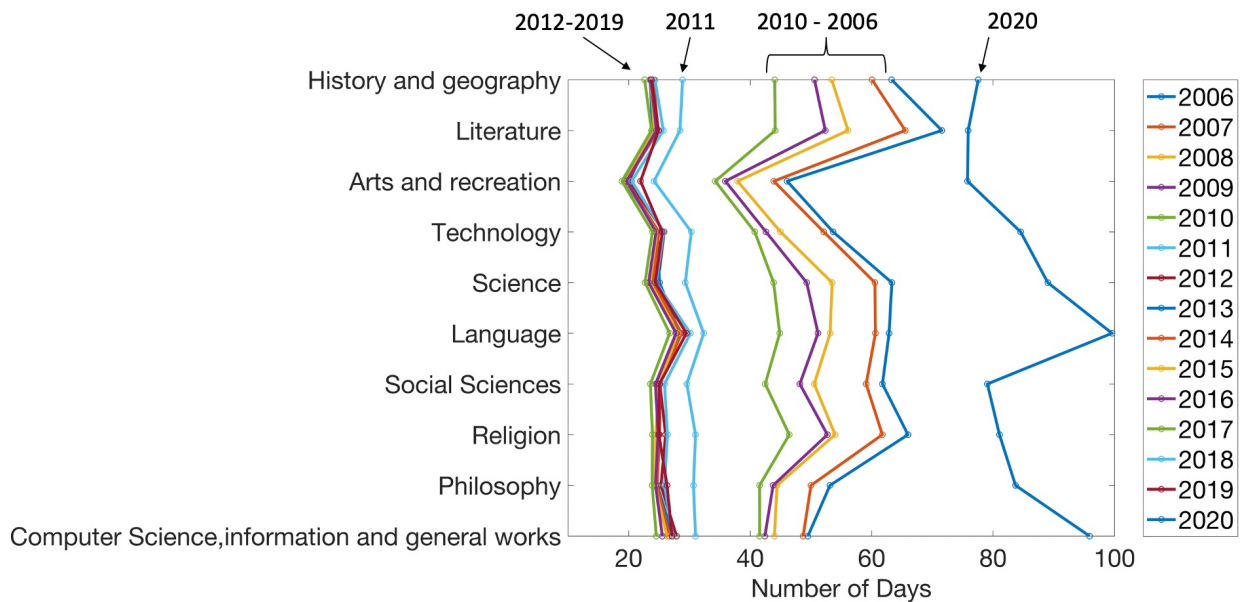


Figure 1 checkout durations for books in different dewey classes from 2006-2020

To look at the relative duration across different dewey classes, I first standardized the duration for each year. In this way, we can have a sense of that within each year, which dewey classes have the longest and shortest checkout durations. From Fig 2, *Arts and recreation* consistently has the shortest checkout durations. *Literature* and *Religion* have the longest checkout durations before 2011, and *language* has the longest checkout durations after 2011.

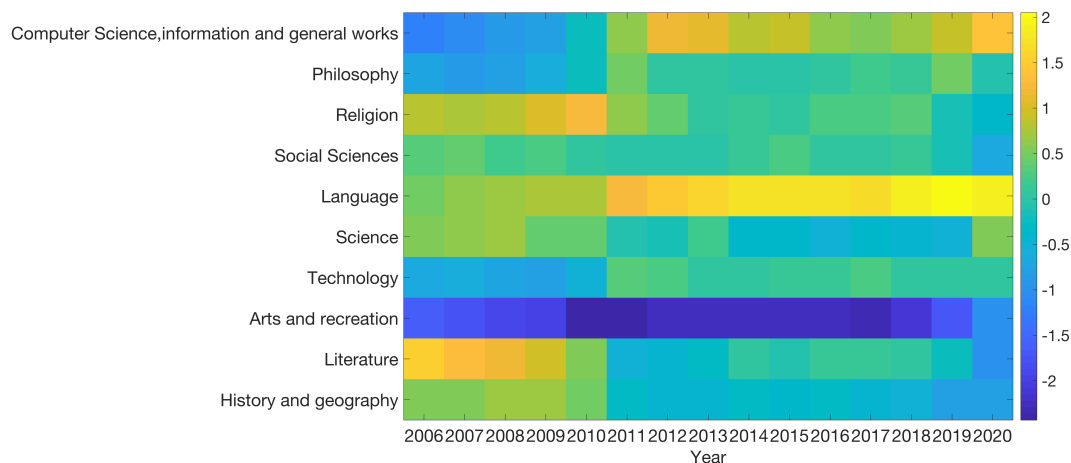


Figure 2 standardized checkout durations for each year from 2006-2020

To look more into the classes with the longest and shortest checkout durations, I looked at the subclasses for *arts and recreation*, *language* and *literature*.

Fig 3 shows a similar plot as Fig 2 but on the subclasses in *arts and recreation* class. From Fig 3, *graphic arts and decorative arts* has the shortest duration across most of the years. Fig 4 shows the number of checkouts for different subclasses. *Graphic arts and decorative arts* again has by far the most checkouts. From Fig 3 and Fig 4, we can conclude that *Graphic arts and decorative arts* contributes the most to the reason why *arts and recreation* has the shortest checkout durations.

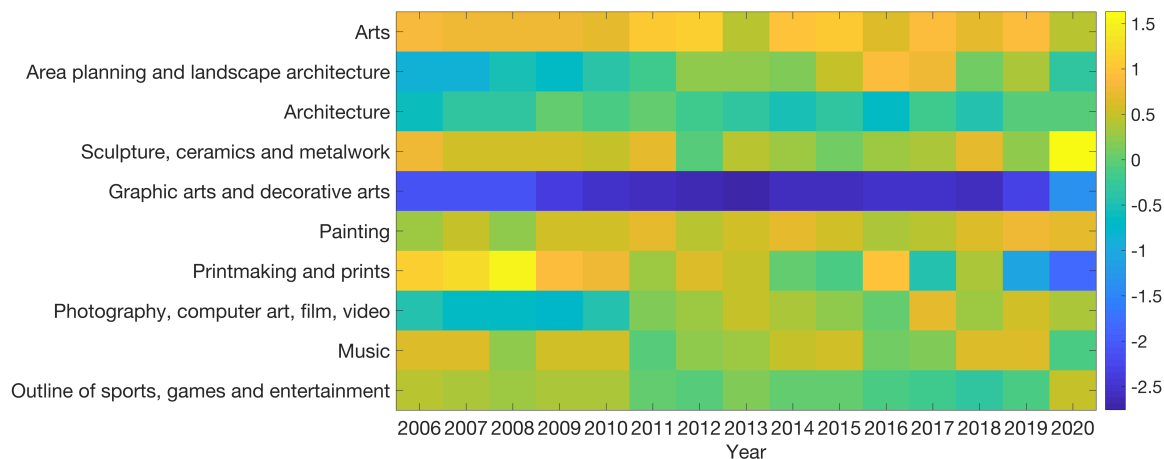


Figure 3 standardized checkout durations for arts and decorative class each year from 2006-2020

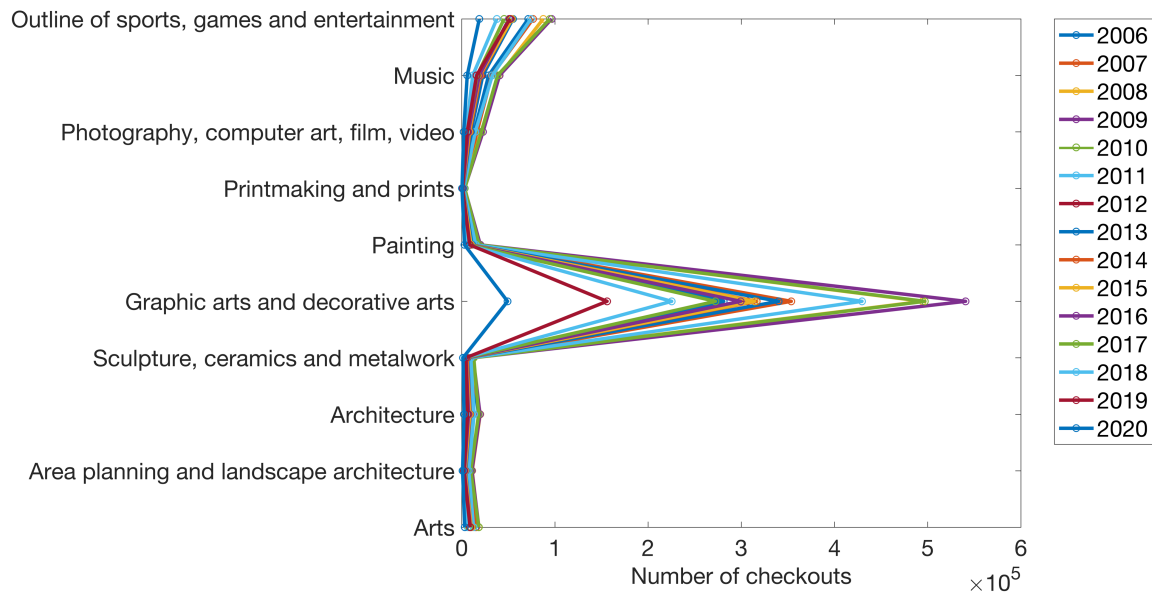


Figure 4 the number of checkouts for subclasses in arts and recreation class from year 2006-2020

Same analysis is done on the *language* and *literature* classes. Our purpose to look at these two classes is to see which subclasses contribute most to the long durations. Unfortunately, the results are not as clear. The pattern I can see is that the *other literatures* has the shortest duration in the *literature* class. For other subclasses, the longest checkout durations seem to be from different subclasses across different years.



Figure 5 standardized checkout durations for subclasses in language class each year from 2006-2020

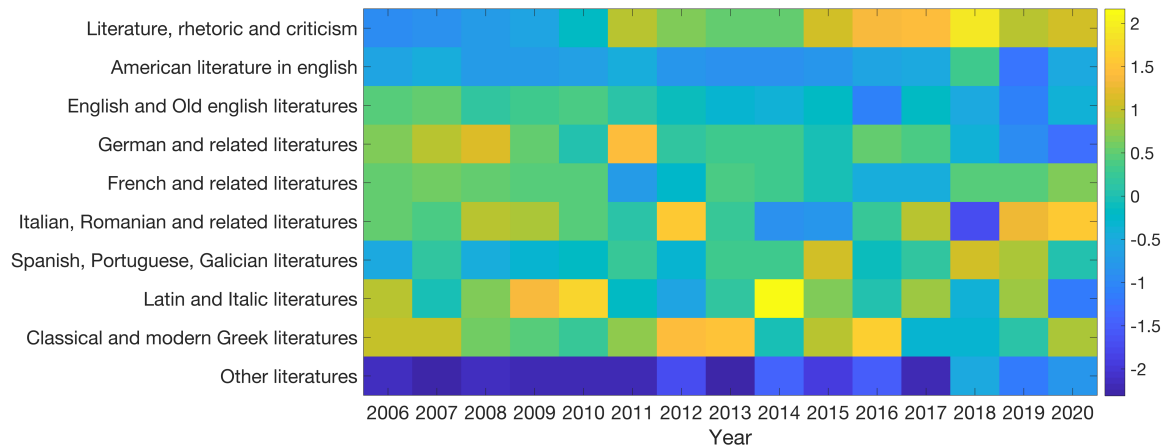


Figure 6 standardized checkout durations for subclasses in literature class each year from 2006-2020

Although there is not a clear pattern from the checkout durations, from looking at the number of checkouts, *English and Old English languages* and *American literature in English* is the most popular subclasses in *language* and *literature* respectively (Fig 7 and Fig 8).

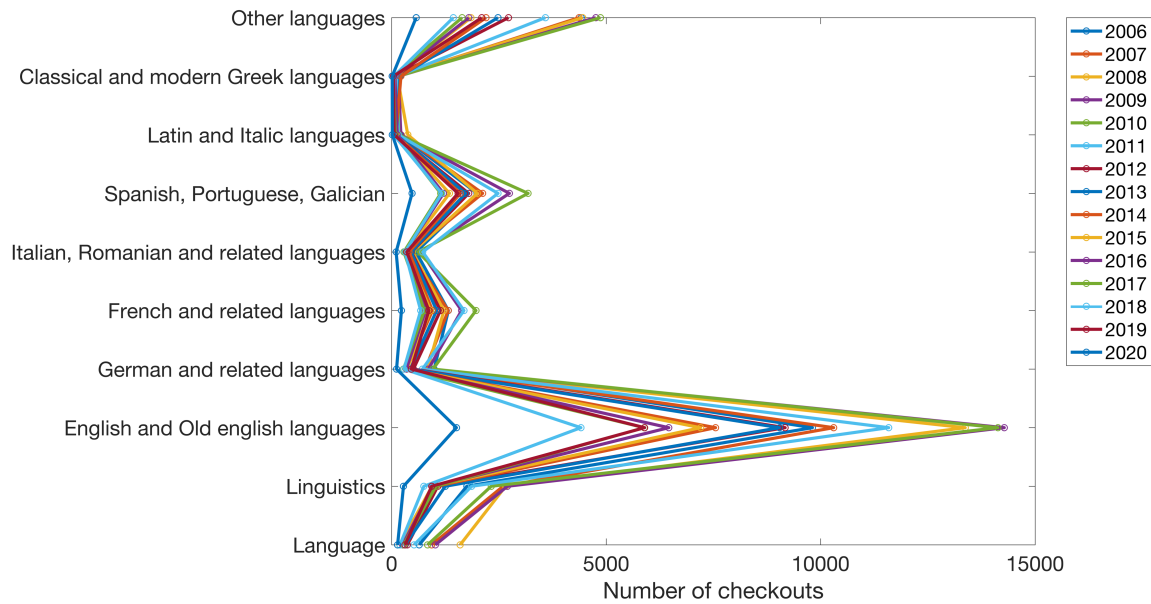


Figure 7 the number of checkouts for subclasses in language class from year 2006-2020

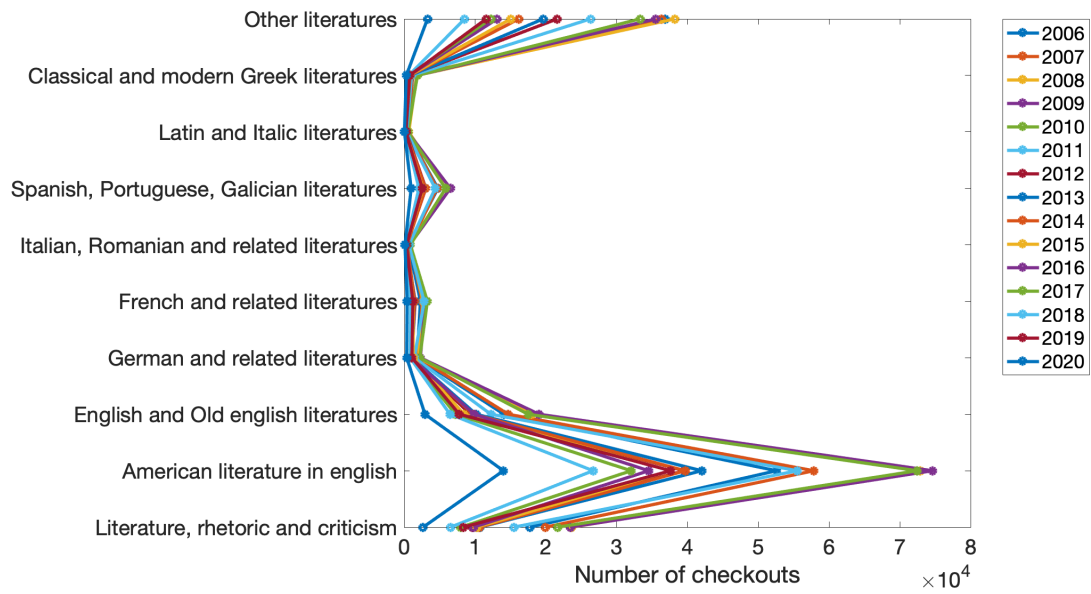


Figure 8 the number of checkouts for subclasses in literature class from year 2006-2020