

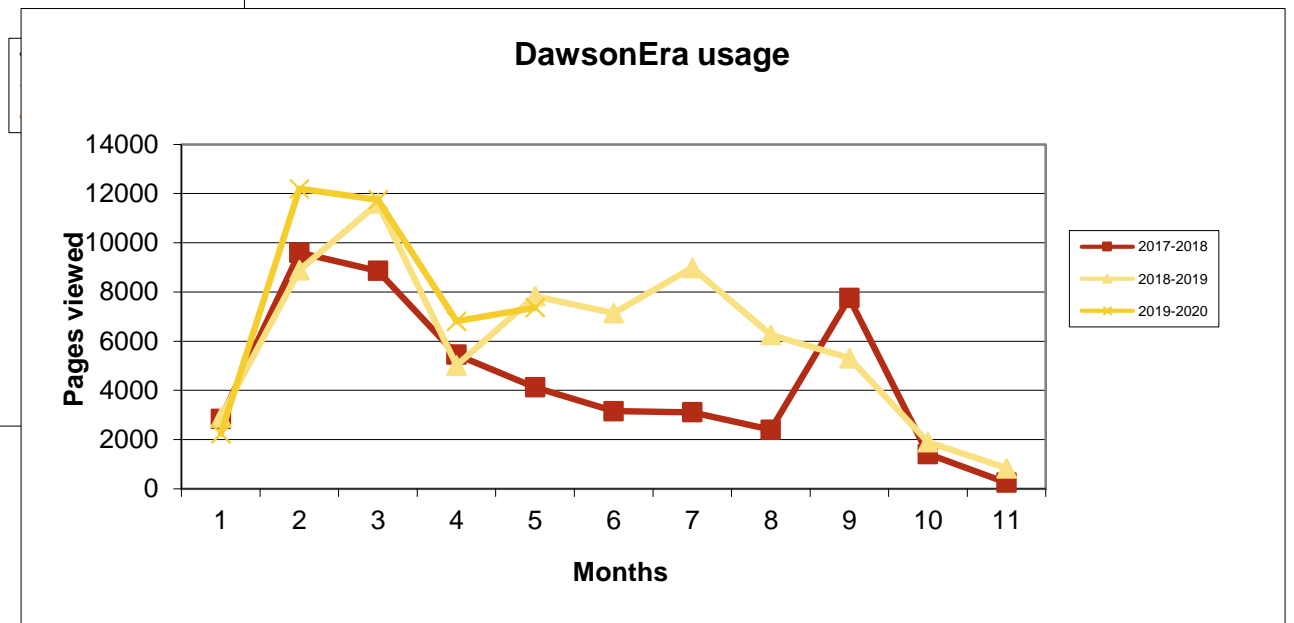
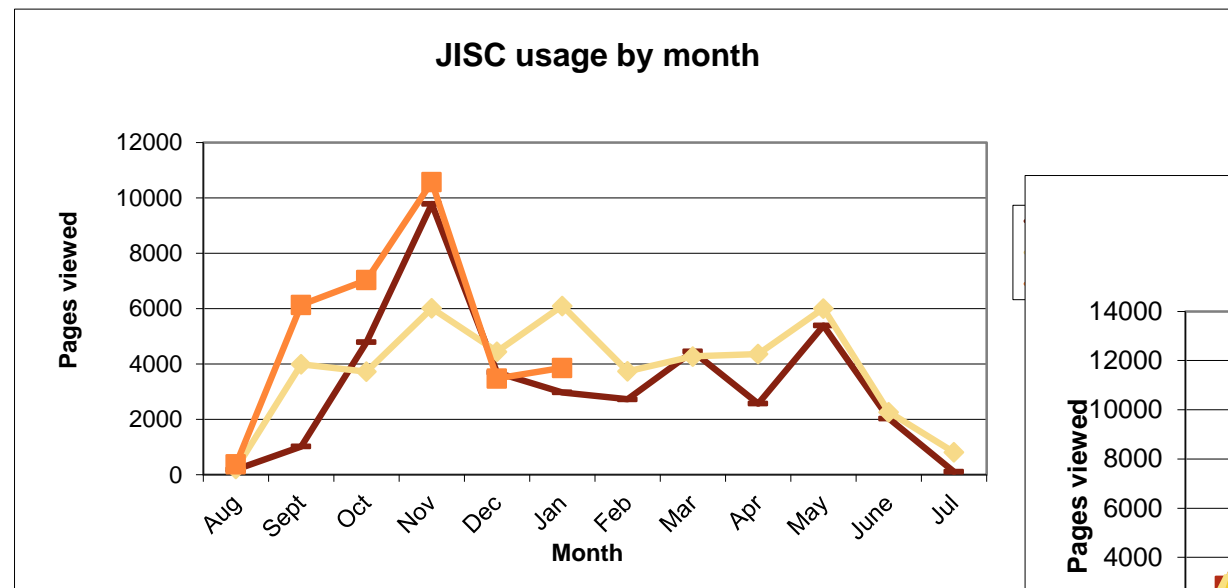


E-book data – what else can we learn?

Sarah Morris – Team Librarian

Evaluating the eBook data

- At present for Performance Review (PR) the statistics are on the basis of pages viewed



More of the story

- **Pages viewed/sessions recorded** is one aspect of the

quantitative data

Feb-20	JISC usage	
Title	eISBN	Total pages
Cache Level 3 Child Care and Education : Early Years Educator	9781471843198	820
Cfe Higher Biology Student Book		568
BTEC National Sport Book 1	9781847766809	532
Btec Level 3 National in It	9781444128512	445
Anatomy and Physiology	9781510435100	355
OCR A Level Biology Student Book 2	9781471827099	323
Edexcel a Level Biology Student Book 2	9781471807381	260
BTEC National Information Technology Practitioners Book 1	9781847766786	251
WJEC/Eduqas A-level Year 2 Biology Student Guide : Variation, Inheritance and Options	9781471859380	249
CACHE Level 3 Early Years Educator : for the Work-Based Learner	9781471808074	234

reports

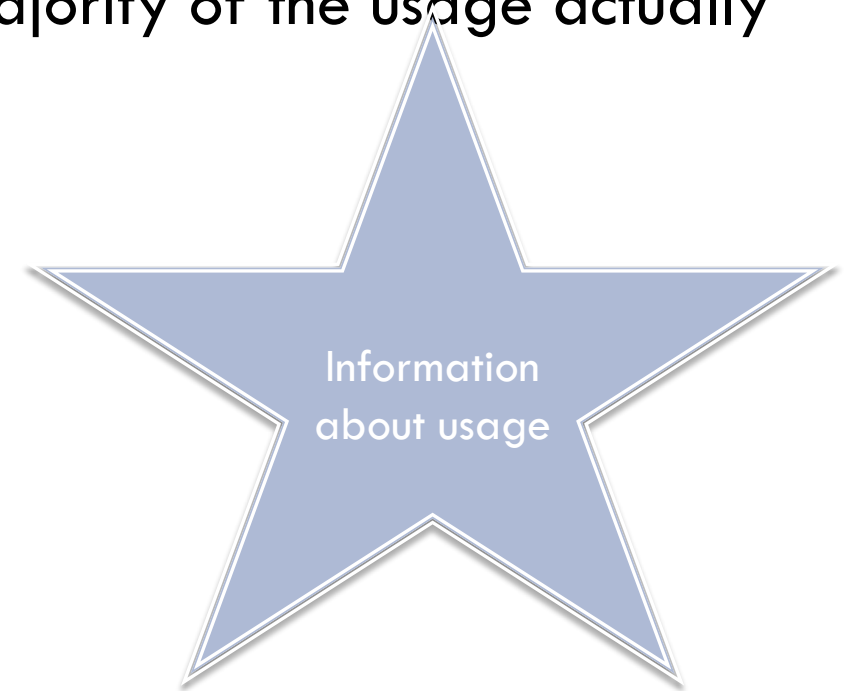
onEra

rence in

DawsonEra usage		Feb-20
Title	eISBN	Total pages
BTEC Level 3 National Sport Book 2	'9781447969723	684
BTEC Level 3 National Public Services Student Book 1	'9781447969655	534
Fundamentals of Anatomy and Physiology	'9781119063735	529
BIOS Instant Notes in Genetics	'9780203079492	510
Genetics, revised edition	'9781908625168	453
Clinical Endocrinology and Diabetes at a Glance	'9781119128724	410
Sickle cell disease	'9781438118123	362
Diabetes For Dummies	'9780470977545	277
Genetics For Dummies	'9780470634462	242
A Complete Guide to the Level 5 Diploma in Education and Training	'9781910391808	233

Investigation

- What did I want to know about how students were using the resources?
 - ▣ How long on average was a user spending in an eBook?
 - ▣ Do we have a lot of repeat users, or is the majority of the usage actually single use?
 - ▣ What was the spread of titles being used?
- What dataset to use?
 - ▣ JISC – most detailed data



Literature search

- Identifying a benchmark
 - ▣ Minimal discussion of the use of eBooks in FE
 - ▣ Most discussion looked at usability issues
 - ▣ Most discussion was looking at HE academic library usage often comparing disciplinary areas

Results so far

Users interaction with eBooks

- Briefly reviewing our usage data it appeared that the time a user spent in an eBook was quite short <5mins
- A paper from 2017 studied the eBook user experience from usage data and user tests of users at Purdue University concluded that 69% reading sessions lasted less than 10 minutes, and 92% sessions lasted less than 30 minutes.
- College data

	September	October	November	December	January		Purdue
Average duration	00:07:34	00:07:40	00:07:27	00:07:27	00:06:50		
Sessions <10mins	86%	85%	78%	87%	89%		69%
Sessions <30 mins	92%	99%	93%	98%	102%		92%

Repeat users or single users

- JISC reports on the source IP address for the device which accesses the ebook
- Established with information from our MITS colleagues the on campus student computer addresses IP ranges
 - ▣ Students using their own devices on the College WiFi would report as the device IP address, not the College WiFi

IP address analysis

□ From September 2019

- the majority of unique IP addresses recorded have only been used once
 - 376 different IP addresses have been used over six months
 - 33 IP addresses have been used in two of the five months
 - IP addresses attributable to our sites at Bournemouth and Lansdowne are the only ones which have been used every month
- After the first month (September) as a percentage the majority of students were using their own devices to access eBooks but onsite access showed the greatest volume

Is this a concern?

What is the spread of titles used?

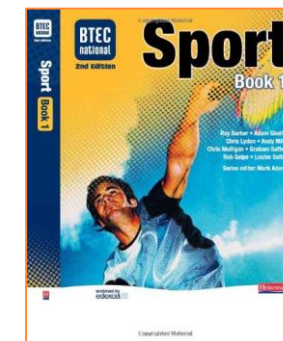
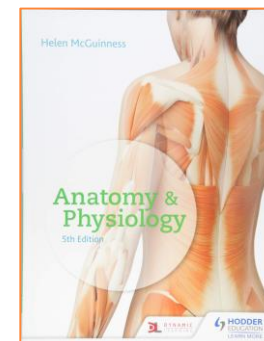
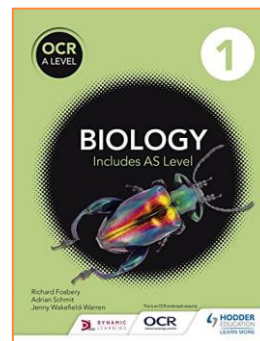


- On the basis of the top 10 titles recorded month by month
 - ▣ What was the number of titles involved, was there variance in the titles from month to month?
 - ▣ Of the titles in use, which were the highest in terms of
 - Pages read
 - Months used

Top 10 results Sept-Jan

- How many different titles appeared in the top 10s?
 - ▣ 28 different titles appeared in the top 10s
 - ▣ Only 1 title was used in all five months
- Which titles cumulatively had the highest number of pages read?
 - ▣ Anatomy and Physiology
 - ▣ BTEC Sport Book 1
- Which titles were used the most consistently
 - ▣ Anatomy and Physiology
 - ▣ BTEC Sport Book 1
 - ▣ OCR Biology

Five months popularity, Sept 19-Jan 20 based on top 10 data only					
Of 28 unique titles					
Number of months used	1	2	3	4	5
Unique titles	15	8	2	2	1



How did I do this?

- Excel - some useful things to know
 - ▣ How to add multiple adjacent cells
 - ▣ Identify and list unique entries
 - ▣ Count the number of times a certain entry is found
 - ▣ How to anchor cells - \$

SUM

- =SUM(range of cells) The cell range can be highlighted using the mouse

In the destination cell enter

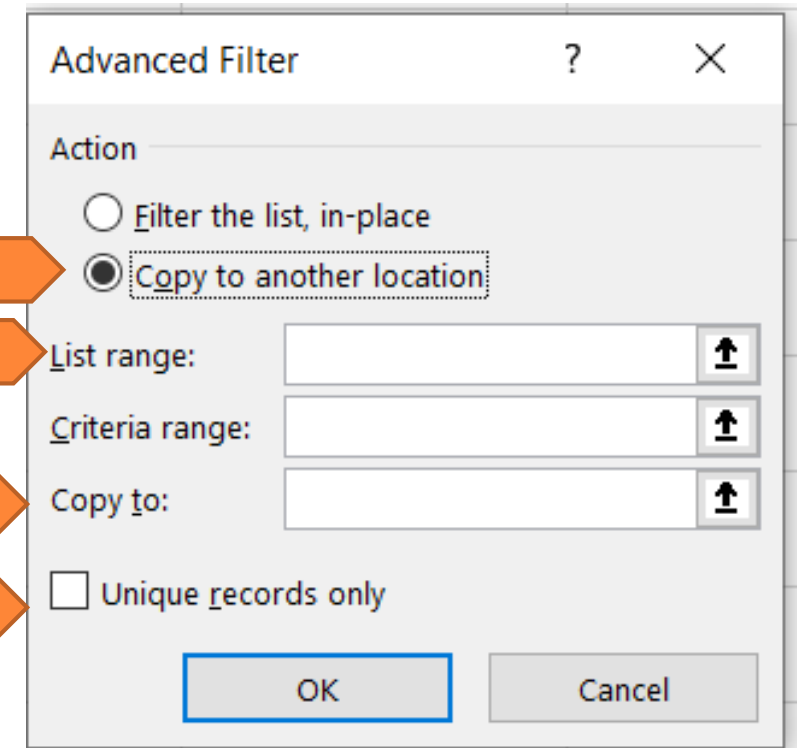
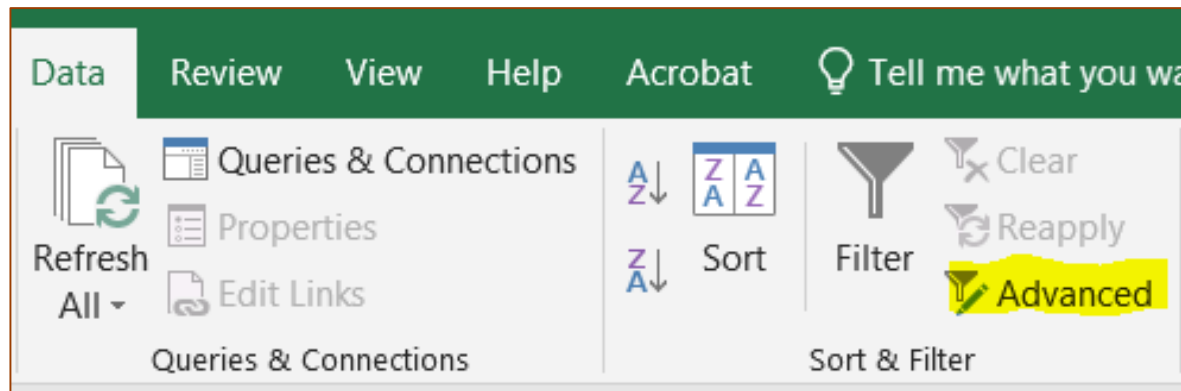
=SUM(C9:c40)

which will add up all the numbers in that column range, much quicker than typing the formula of =(c9+c10+c11....c40)

- There is also a =SUMIF function which can be used to carry out mathematical operations on cells which meet a users specified criteria

Identify and list unique entries

- This is an advanced filter function on the data menu



COUNTIF

- This function will search a user specified range of cells, and count the number of occurrences
 - ▣ The formula is typed in a destination cell for the calculation
 - =COUNTIF(start of range:end of range, criteria)
 - The criteria can make use of the normal mathematical functions =, <, > or in combination as long as it is enclosed in quotations
 - =countif(C9:c40, "=123.456.67.9")
 - It is possible to use a cell reference for the criteria
 - =countif(c9:c40, b1) There are no quotation marks when it is a cell reference

Anchoring cells - \$

- The ability to anchor certain cells in a formula can prove very useful, especially when copying and pasting formulae around a spreadsheet.

$$=(c4/\$c\$6)$$

This would always divide whatever value was in the first cell by the contents of cell c6

- The user can choose to anchor just the column or the row as appropriate to the layout of the data

Going forwards

□ The analysis so far shows

Consistency appears to be a hurdle, students will try eBooks, but not adopt them

- ▣ Is this data broadly consistent with what can be established from other similar institutions?
- ▣ Is this data broadly consistent with usage from other providers?
 - DawsonEra data is more limited and cannot provide the same level of detail.

Common fields	Date	Duration/minutes	User	Title	Publisher	eISBN	Total pages read	LCC	Dewey
			as in library	of book					

- Possibly looking at an available proxy like the “operating system data” in the **users** field for the IP address. If it is usable, then February data is 96 unique users.

Goodness

□ In what context?

▣ Data integrity

Limitations:


- IP address usage is not directly attributable to an individual student as some may have more than one device, or have changed their device over the analysis period.
- Staff usage is automatically included in the statistics because individual logins are not recorded.

▣ There is little comparable data

- External validation is on volume (pages viewed etc), is that quantity over quality

▣ Does it show value for money?

- Different measures – cost per page view, only really the end of the academic year



As professionals think about the narrative, you may have to report a single figure, but have a story as well

References

- Zhang, T., Niu, X. and Promann, M., 2017. Assessing the User Experience of e-books in academic libraries. *College and Research Libraries*. 78 (5), 578-601.
- Excel tips available from lots of websites – and the online help.
 - ▣ Exceljet