Hiding in Plain Sight

Identification and Analysis of Anomalous Files

Presented by: Ramece Cave

What You Will Learn

- Common Tools
- Limitations in Tools
- Analyzing Files
- Identifying File Structure
- Caveats

Bio

- Began working in information security in the Internet Abuse department at UUNET in 1999. Over the past 10 years have been focusing on forensics, reverse engineering, and malware analysis; in various incident response and SOC positions. Currently work as a Security/Malware Engineer at CompuCom, developing tools and techniques for malware analysis and enhancing incident response measures.

The Problem

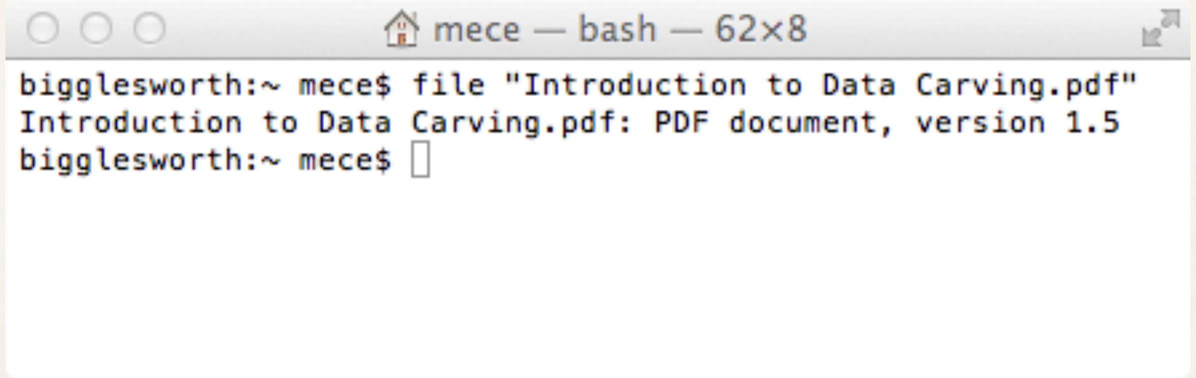
- Thousands of businesses and organizations are often the unwilling and unknowing distributers of malware or related components

- The associated files appear normal or harmless, using different file extensions to avoid detection.

- Anti-Virus does not typically detect non-executable formats.

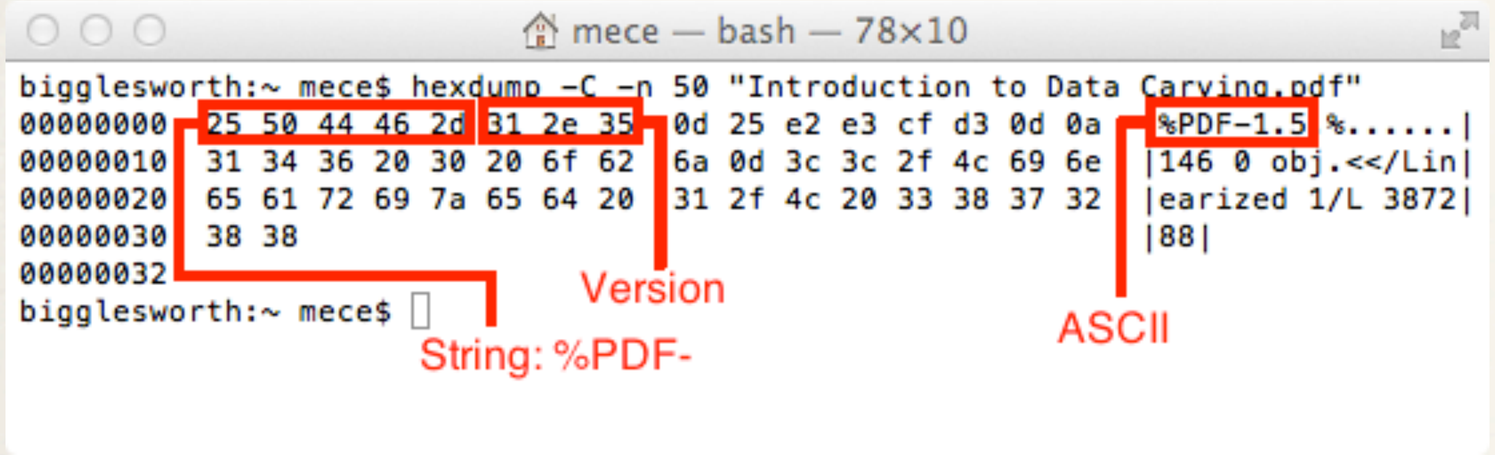
Common Tools

- The most popular and common tool used for identifying file types is the command-line based tool is “File”. File uses signatures (similar to an IDS) to identify specific characteristics in a file and determine its type of program association.
-



```
bigglesworth:~ mece$ file "Introduction to Data Carving.pdf"
Introduction to Data Carving.pdf: PDF document, version 1.5
bigglesworth:~ mece$ 
```

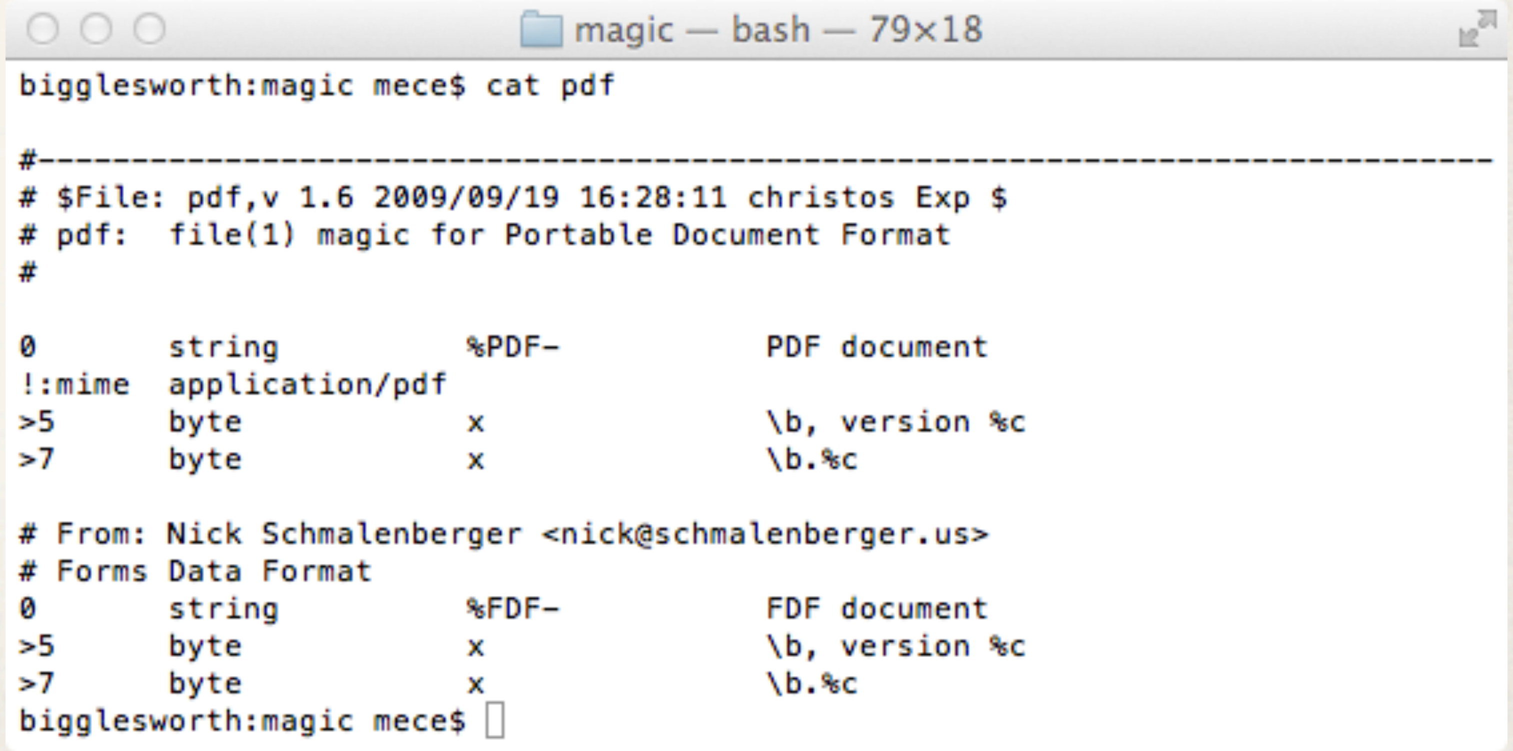
Common Tools (Cont)



```
biggleswerth:~ mece$ hexdump -C -n 50 "Introduction to Data Carving.pdf"
00000000  25 50 44 46 2d 31 2e 35  0d 25 e2 e3 cf d3 0d 0a  |%PDF-1.5 %......|
00000010  31 34 36 20 30 20 6f 62  6a 0d 3c 3c 2f 4c 69 6e  |146 0 obj.<</Lin|
00000020  65 61 72 69 7a 65 64 20  31 2f 4c 20 33 38 37 32  |earized 1/L 3872|
00000030  38 38                                             |88|
00000032
biggleswerth:~ mece$ 
```

Hex dump of the first 50 bytes of the file.

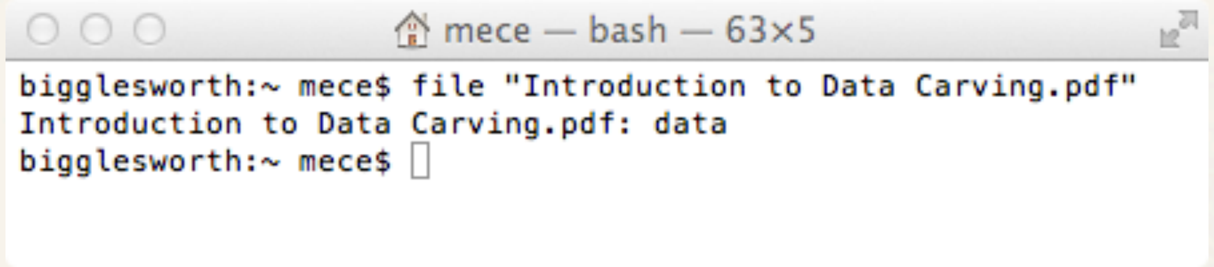
PDF "magic" file located in : /usr/share/file/magic



```
biggleswerth:magic mece$ cat pdf

#------------------------------------------------------------------------------
# $File: pdf,v 1.6 2009/09/19 16:28:11 christos Exp $
# pdf:  file(1) magic for Portable Document Format
#

0       string          %PDF-           PDF document
!:mime  application/pdf
>5      byte            x               \b, version %c
>7      byte            x               \b.%c

# From: Nick Schmalenberger <nick@schmalenberger.us>
# Forms Data Format
0       string          %FDF-           FDF document
>5      byte            x               \b, version %c
>7      byte            x               \b.%c
biggleswerth:magic mece$ 
```

Limitations

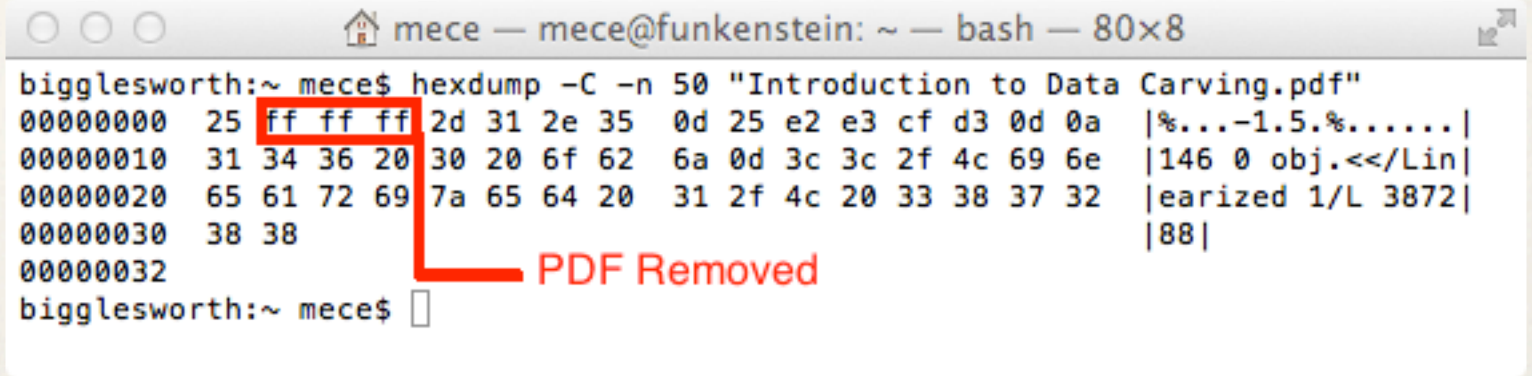
Altered PDF document is no longer identifiable



```
bigglesworth:~ mece$ file "Introduction to Data Carving.pdf"
Introduction to Data Carving.pdf: data
bigglesworth:~ mece$ 
```

- File is a great tool, but when it cannot identify a file, it simply returns "data" as the type. Data ambiguously, indicates the file has not previously been identified.

PDF [\x50\x44\x46] removed

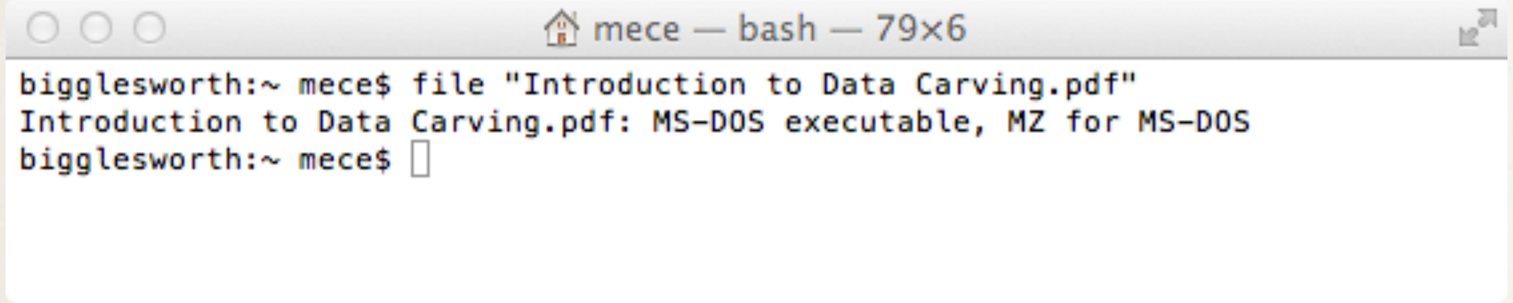


```
bigglesworth:~ mece$ hexdump -C -n 50 "Introduction to Data Carving.pdf"
00000000  25 ff ff ff 2d 31 2e 35  0d 25 e2 e3 cf d3 0d 0a  |%...-1.5.%......|
00000010  31 34 36 20 30 20 6f 62  6a 0d 3c 3c 2f 4c 69 6e  |146 0 obj.<</Lin|
00000020  65 61 72 69 7a 65 64 20  31 2f 4c 20 33 38 37 32  |earized 1/L 3872|
00000030  38 38                                             |88|
00000032
bigglesworth:~ mece$ 
```

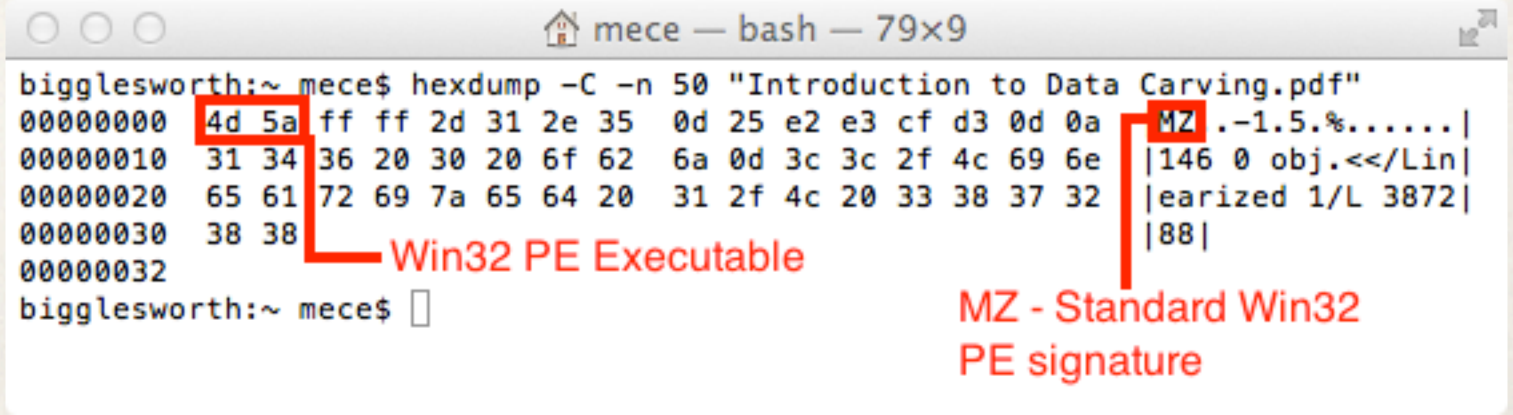
PDF Removed

Analyzing Files

- Trust but verify File results.
- Look for patterns.
- Working in bulk can be a big help.



```
bigglesworth:~ mece$ file "Introduction to Data Carving.pdf"
Introduction to Data Carving.pdf: MS-DOS executable, MZ for MS-DOS
bigglesworth:~ mece$ 
```



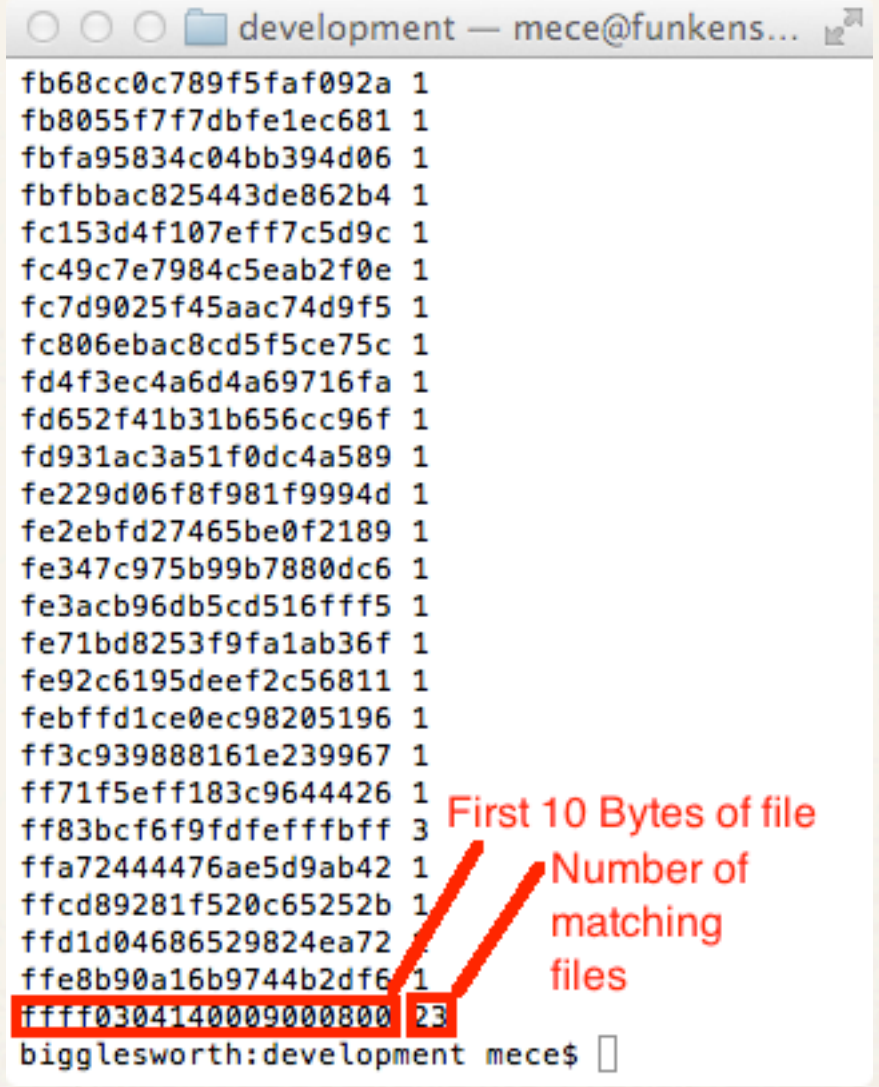
```
bigglesworth:~ mece$ hexdump -C -n 50 "Introduction to Data Carving.pdf"
00000000  4d 5a ff ff 2d 31 2e 35  0d 25 e2 e3 cf d3 0d 0a  |MZ..-1.5.%......|
00000010  31 34 36 20 30 20 6f 62  6a 0d 3c 3c 2f 4c 69 6e  |146 0 obj.<</Lin|
00000020  65 61 72 69 7a 65 64 20  31 2f 4c 20 33 38 37 32  |earized 1/L 3872|
00000030  38 38                                             |88|
00000032
bigglesworth:~ mece$ 
```

Win32 PE Executable

MZ - Standard Win32
PE signature

Analyzing Files (cont)

Bulk analysis of all files in a given directory



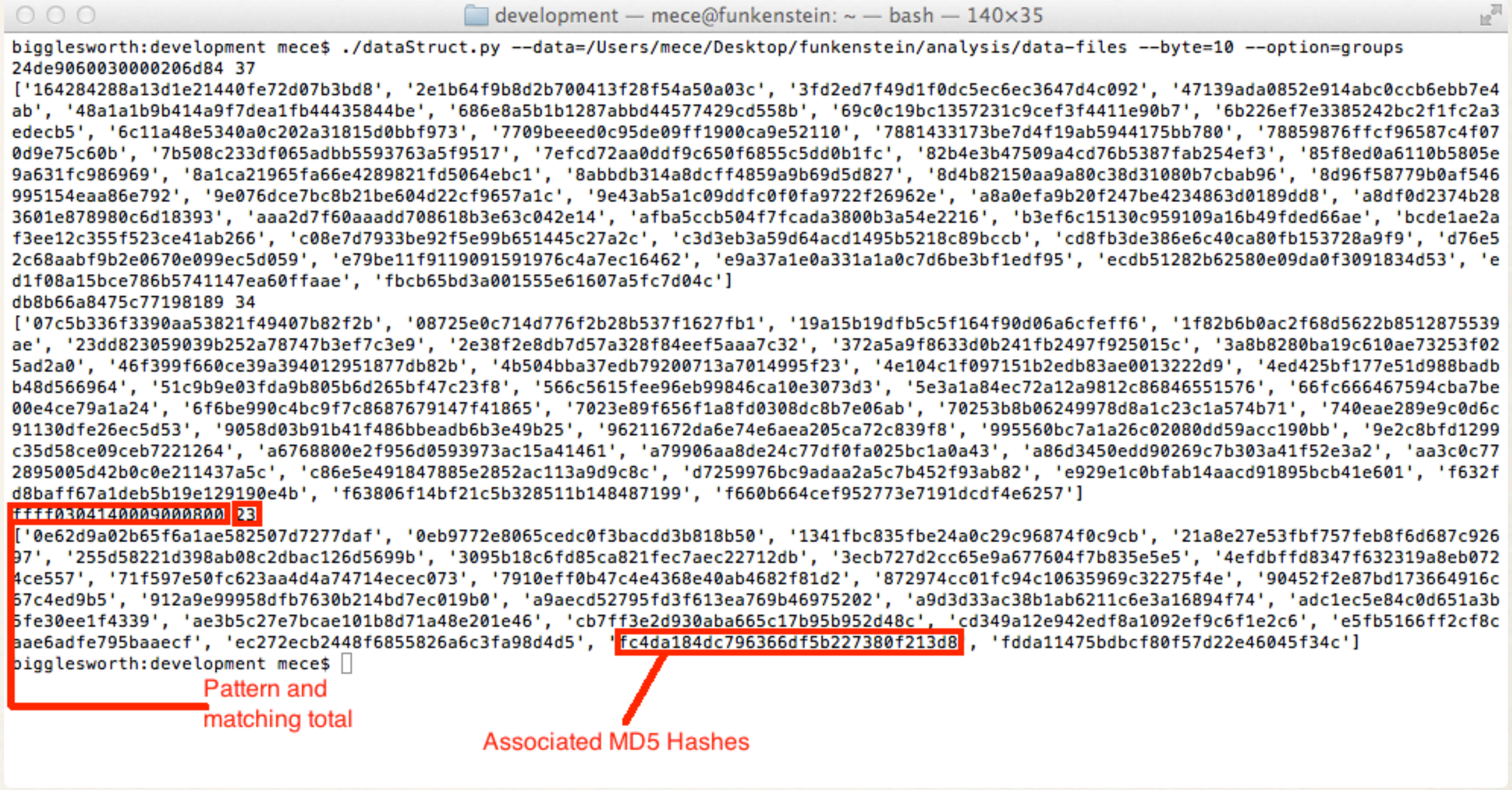
```
fb68cc0c789f5faf092a 1
fb8055f7f7dbfe1ec681 1
fbfa95834c04bb394d06 1
fbfbbac825443de862b4 1
fc153d4f107eff7c5d9c 1
fc49c7e7984c5eab2f0e 1
fc7d9025f45aac74d9f5 1
fc806ebac8cd5f5ce75c 1
fd4f3ec4a6d4a69716fa 1
fd652f41b31b656cc96f 1
fd931ac3a51f0dc4a589 1
fe229d06f8f981f9994d 1
fe2ebfd27465be0f2189 1
fe347c975b99b7880dc6 1
fe3acb96db5cd516fff5 1
fe71bd8253f9fa1ab36f 1
fe92c6195deef2c56811 1
febffd1ce0ec98205196 1
ff3c939888161e239967 1
ff71f5eff183c9644426 1
ff83bcf6f9fdfefffbff 3
ffa72444476ae5d9ab42 1
ffcd89281f520c65252b 1
ffd1d04686529824ea72 1
ffe8b90a16b9744b2df6 1
ffff0304140009000800 23
bigglesworth:development mece$ 
```

First 10 Bytes of file

Number of matching files

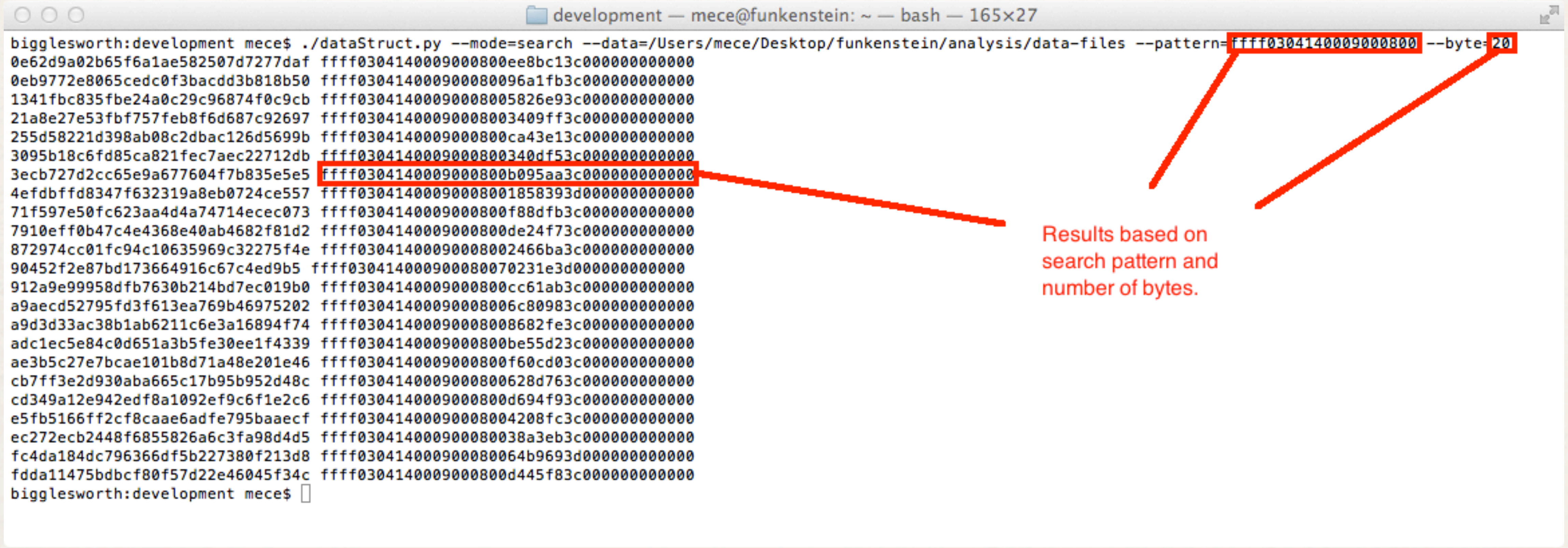
Analyzing Files (cont)

Grouped results based on top patterns



```
biggles worth:development mece$ ./dataStruct.py --data=/Users/mece/Desktop/funkenstein/analysis/data-files --byte=10 --option=groups
24de9060030000206d84 37
['164284288a13d1e21440fe72d07b3bd8', '2e1b64f9b8d2b700413f28f54a50a03c', '3fd2ed7f49d1f0dc5ec6ec3647d4c092', '47139ada0852e914abc0ccb6ebb7e4
ab', '48a1a1b9b414a9f7dea1fb44435844be', '686e8a5b1b1287abbd44577429cd558b', '69c0c19bc1357231c9cef3f4411e90b7', '6b226ef7e3385242bc2f1fc2a3
edecb5', '6c11a48e5340a0c202a31815d0bbf973', '7709beeed0c95de09ff1900ca9e52110', '7881433173be7d4f19ab5944175bb780', '78859876ffcf96587c4f07
0d9e75c60b', '7b508c233df065adbb5593763a5f9517', '7efcd72aa0ddf9c650f6855c5dd0b1fc', '82b4e3b47509a4cd76b5387fab254ef3', '85f8ed0a6110b5805e
9a631fc986969', '8a1ca21965fa66e4289821fd5064ebc1', '8abbdb314a8dcff4859a9b69d5d827', '8d4b82150aa9a80c38d31080b7cbab96', '8d96f58779b0af546
995154eaa86e792', '9e076dce7bc8b21be604d22cf9657a1c', '9e43ab5a1c09ddfc0f0fa9722f26962e', 'a8a0efa9b20f247be4234863d0189dd8', 'a8df0d2374b28
3601e878980c6d18393', 'aaa2d7f60aaadd708618b3e63c042e14', 'afba5ccb504f7fcada3800b3a54e2216', 'b3ef6c15130c959109a16b49fded66ae', 'bcde1ae2a
f3ee12c355f523ce41ab266', 'c08e7d7933be92f5e99b651445c27a2c', 'c3d3eb3a59d64acd1495b5218c89bccb', 'cd8fb3de386e6c40ca80fb153728a9f9', 'd76e5
2c68aabf9b2e0670e099ec5d059', 'e79be11f9119091591976c4a7ec16462', 'e9a37a1e0a331a1a0c7d6be3bf1edf95', 'ecdb51282b62580e09da0f3091834d53', 'e
d1f08a15bce786b5741147ea60ffaae', 'fbcb65bd3a001555e61607a5fc7d04c']
db8b66a8475c77198189 34
['07c5b336f3390aa53821f49407b82f2b', '08725e0c714d776f2b28b537f1627fb1', '19a15b19dfb5c5f164f90d06a6cfeff6', '1f82b6b0ac2f68d5622b8512875539
ae', '23dd823059039b252a78747b3ef7c3e9', '2e38f2e8db7d57a328f84eef5aaa7c32', '372a5a9f8633d0b241fb2497f925015c', '3a8b8280ba19c610ae73253f02
5ad2a0', '46f399f660ce39a394012951877db82b', '4b504bba37edb79200713a7014995f23', '4e104c1f097151b2edb83ae0013222d9', '4ed425bf177e51d988badb
b48d566964', '51c9b9e03fda9b805b6d265bf47c23f8', '566c5615fee96eb99846ca10e3073d3', '5e3a1a84ec72a12a9812c86846551576', '66fc666467594cba7be
00e4ce79a1a24', '6f6be990c4bc9f7c8687679147f41865', '7023e89f656f1a8fd0308dc8b7e06ab', '70253b8b06249978d8a1c23c1a574b71', '740eae289e9c0d6c
91130dfe26ec5d53', '9058d03b91b41f486bbeadb6b3e49b25', '96211672da6e74e6aea205ca72c839f8', '995560bc7a1a26c02080dd59acc190bb', '9e2c8bfd1299
c35d58ce09ceb7221264', 'a6768800e2f956d0593973ac15a41461', 'a79906aa8de24c77df0fa025bc1a0a43', 'a86d3450edd90269c7b303a41f52e3a2', 'aa3c0c77
2895005d42b0c0e211437a5c', 'c86e5e491847885e2852ac113a9d9c8c', 'd7259976bc9adaa2a5c7b452f93ab82', 'e929e1c0bfab14aacd91895bcb41e601', 'f632f
d8baff67a1deb5b19e129190e4b', 'f63806f14bf21c5b328511b148487199', 'f660b664cef952773e7191dcdf4e6257']
ffff0304140009000800 23
['0e62d9a02b65f6a1ae582507d7277daf', '0eb9772e8065cedc0f3bacdd3b818b50', '1341fbc835fbe24a0c29c96874f0c9cb', '21a8e27e53fbf757feb8f6d687c926
97', '255d58221d398ab08c2dbac126d5699b', '3095b18c6fd85ca821fec7aec22712db', '3ecb727d2cc65e9a677604f7b835e5e5', '4efdbffd8347f632319a8eb072
4ce557', '71f597e50fc623aa4d4a74714ecec073', '7910eff0b47c4e4368e40ab4682f81d2', '872974cc01fc94c10635969c32275f4e', '90452f2e87bd173664916c
57c4ed9b5', '912a9e99958dfb7630b214bd7ec019b0', 'a9aecd52795fd3f613ea769b46975202', 'a9d3d33ac38b1ab6211c6e3a16894f74', 'adc1ec5e84c0d651a3b
5fe30ee1f4339', 'ae3b5c27e7bcae101b8d71a48e201e46', 'cb7ff3e2d930aba665c17b95b952d48c', 'cd349a12e942edf8a1092ef9c6f1e2c6', 'e5fb5166ff2cf8c
aae6adfe795baaecf', 'ec272ecb2448f6855826a6c3fa98d4d5', 'fc4da184dc796366df5b227380f213d8', 'fdda11475bdbcf80f57d22e46045f34c']
biggles worth:development mece$
```

Analyzing Files (cont)



```
bigglesworth:development mece$ ./dataStruct.py --mode=search --data=/Users/mece/Desktop/funkenstein/analysis/data-files --pattern=ffff0304140009000800 --byte=20
0e62d9a02b65f6a1ae582507d7277daf ffff0304140009000800ee8bc13c000000000000
0eb9772e8065cedc0f3bacdd3b818b50 ffff030414000900080096a1fb3c000000000000
1341fbc835fbe24a0c29c96874f0c9cb ffff03041400090008005826e93c000000000000
21a8e27e53fbf757feb8f6d687c92697 ffff03041400090008003409ff3c000000000000
255d58221d398ab08c2dbac126d5699b ffff0304140009000800ca43e13c000000000000
3095b18c6fd85ca821fec7aec22712db ffff0304140009000800340df53c000000000000
3ecb727d2cc65e9a677604f7b835e5e5 ffff0304140009000800b095aa3c000000000000
4efdbffd8347f632319a8eb0724ce557 ffff03041400090008001858393d000000000000
71f597e50fc623aa4d4a74714ecec073 ffff0304140009000800f88dfb3c000000000000
7910eff0b47c4e4368e40ab4682f81d2 ffff0304140009000800de24f73c000000000000
872974cc01fc94c10635969c32275f4e ffff03041400090008002466ba3c000000000000
90452f2e87bd173664916c67c4ed9b5 ffff030414000900080070231e3d000000000000
912a9e99958dfb7630b214bd7ec019b0 ffff0304140009000800cc61ab3c000000000000
a9aecd52795fd3f613ea769b46975202 ffff03041400090008006c80983c000000000000
a9d3d33ac38b1ab6211c6e3a16894f74 ffff03041400090008008682fe3c000000000000
adc1ec5e84c0d651a3b5fe30ee1f4339 ffff0304140009000800be55d23c000000000000
ae3b5c27e7bcae101b8d71a48e201e46 ffff0304140009000800f60cd03c000000000000
cb7ff3e2d930aba665c17b95b952d48c ffff0304140009000800628d763c000000000000
cd349a12e942edf8a1092ef9c6f1e2c6 ffff0304140009000800d694f93c000000000000
e5fb5166ff2cf8caae6adfe795baaecf ffff03041400090008004208fc3c000000000000
ec272ecb2448f6855826a6c3fa98d4d5 ffff030414000900080038a3eb3c000000000000
fc4da184dc796366df5b227380f213d8 ffff030414000900080064b9693d000000000000
fdda11475bdbcf80f57d22e46045f34c ffff0304140009000800d445f83c000000000000
bigglesworth:development mece$ 
```

Results based on search pattern and number of bytes.

Focusing on a group of files.

Identifying File Structure

- Think Simple: repeating patterns indicates structure.
- Arbitrary values or numbers between patterns may be significant.
- Look for ASCII strings within the file.
- Remember: Big vs Little ENDIAN.

Caveats

- The analysis process may take long time.
- Analyzing multiple samples may yield better results.
- 100% conclusive identification without a the originating file may not be possible.
- Unidentified file structures may be based on or a derivative of another well known file type.

Contact Info

- Website: http://www.n00dle.org
- Twitter: @feedbrain
- E-Mail: rrcave@n00dle.org
- **CompuCom Info**
- Twitter: @compucomgsirtres
- Blog: http://compucommssresearchgroup.blogspot.com
- RSS Feed: http://www.google.com/reader/shared/02995287658117904101
- E-Mail: rcave@compucom.com