

Projeto:Parte 2

Criar Arquivo Invertido

Estrutura do dado

```
7 public class TermData {  
8     //starts with zero and increment for each 3 lines separation in csv document  
9     private int position;  
10    private int frequency;  
11    private String documentID;  
12  
13    public TermData(int position, int frequency, String documentID) {  
14        this.position = position;  
15        this.frequency = frequency;  
16        this.documentID = documentID;  
17    }  
18
```

Estrutura do dado

```
5  
6 public class IndexRow {  
7     private String word;  
8     private List<TermData> posting;  
9     private int position;  
10  
11 public IndexRow(String word) {  
12     this.word = word;  
13     this.posting = new ArrayList<TermData>();  
14     this.position = 0;  
15 }  
16
```

```
9 public class InvertedIndex {  
10     private Vector<IndexRow> indexRows;  
11  
12     public InvertedIndex() {  
13         this.indexRows = new Vector<>();  
14         setIndexRowDefaultElements();  
15     }  
16 }
```

```
@Override
1 public String toString()
2 {
3     StringBuilder sb = new StringBuilder();
4
5
6     for (IndexRow indexRow : indexRows) {
7         sb.append(indexRow.getWord());
8         sb.append(";");
9
10        sb.append(indexRow.getPosting().size());
11
12        for (TermData posting : indexRow.getPosting()) {
13            sb.append(posting.toString());
14            sb.append(",");
15        }
16        sb.setLength(sb.length() - 1);
17        sb.append("\n");
18    }
19
20    return sb.toString();
21 }
```

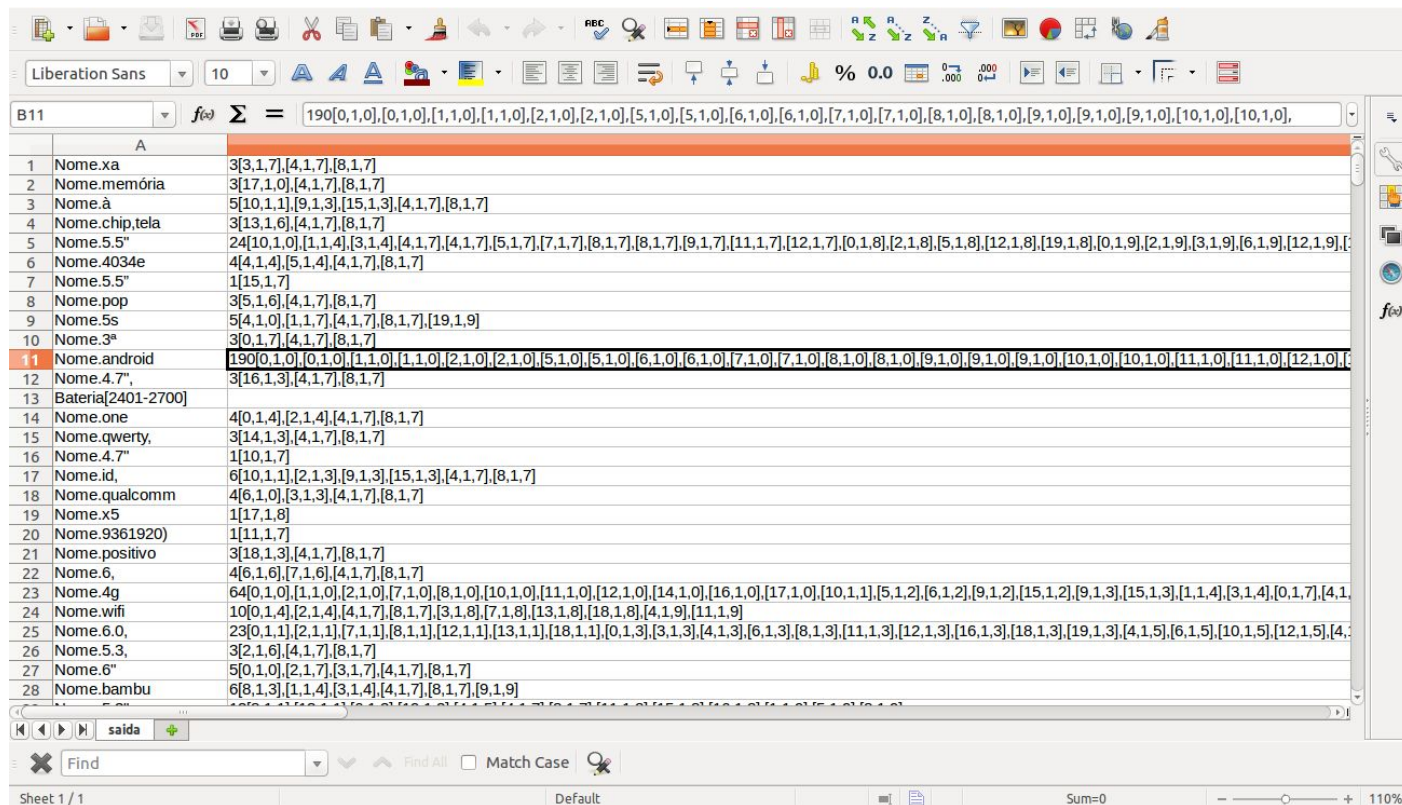
```
33 public String compressedToString(int tamCSVs[])
34 {
35     StringBuilder sb = new StringBuilder();
36     |
37     for (IndexRow indexRow : indexRows) {
38         sb.append(indexRow.getWord());
39         sb.append(";");
40
41         sb.append(indexRow.getPosting().size());
42
43         for (TermData posting : indexRow.getPosting()) {
44             sb.append(posting.compressedToString(tamCSVs));
45             sb.append(",");
46         }
47         sb.setLength(sb.length() - 1);
48         sb.append("\n");
49     }
50
51     return sb.toString();
52 }
53 }
```

```
51
52 public String compressedToString(int tamCSVs[])
53 {
54     int valor = this.position;
55     int doc = Integer.parseInt(getId());
56     for (int i = 0; i < doc; i++)
57     {
58         valor += tamCSVs[i];
59     }
60     return "[" + valor + "," + this.frequency + "]";
61 }
```

Lista invertida

FrequênciaTotal[posiçãoCSV, frequênciaLocal, NomeCsv]

45,7 kB



The screenshot shows a spreadsheet application with a toolbar at the top. The formula bar displays the formula $f_{\text{CSV}} \Sigma =$ followed by a long array of frequency data. The spreadsheet contains a list of items in column A and their corresponding frequency data in column B. Row 11 is highlighted in orange.

	A	B
1	Nome.xa	3[3,1,7],[4,1,7],[8,1,7]
2	Nome.memória	3[17,1,0],[4,1,7],[8,1,7]
3	Nome.à	5[10,1,1],[9,1,3],[15,1,3],[4,1,7],[8,1,7]
4	Nome.chip,tela	3[13,1,6],[4,1,7],[8,1,7]
5	Nome.5.5"	24[10,1,0],[1,1,4],[3,1,4],[4,1,7],[4,1,7],[5,1,7],[7,1,7],[8,1,7],[8,1,7],[9,1,7],[11,1,7],[12,1,7],[0,1,8],[2,1,8],[5,1,8],[12,1,8],[19,1,8],[0,1,9],[2,1,9],[3,1,9],[6,1,9],[12,1,9],[
6	Nome.4034e	4[4,1,4],[5,1,4],[4,1,7],[8,1,7]
7	Nome.5.5"	1[15,1,7]
8	Nome.pop	3[5,1,6],[4,1,7],[8,1,7]
9	Nome.5s	5[4,1,0],[1,1,7],[4,1,7],[8,1,7],[19,1,9]
10	Nome.3ª	3[0,1,7],[4,1,7],[8,1,7]
11	Nome.android	190[0,1,0],[0,1,0],[1,1,0],[2,1,0],[2,1,0],[5,1,0],[5,1,0],[6,1,0],[6,1,0],[7,1,0],[7,1,0],[8,1,0],[8,1,0],[9,1,0],[9,1,0],[9,1,0],[10,1,0],[10,1,0],[11,1,0],[11,1,0],[12,1,0],[
12	Nome.4.7"	3[16,1,3],[4,1,7],[8,1,7]
13	Bateria[2401-2700]	
14	Nome.one	4[0,1,4],[2,1,4],[4,1,7],[8,1,7]
15	Nome.qwerty,	3[14,1,3],[4,1,7],[8,1,7]
16	Nome.4.7"	1[10,1,7]
17	Nome.id,	6[10,1,1],[2,1,3],[9,1,3],[15,1,3],[4,1,7],[8,1,7]
18	Nome.qualcomm	4[6,1,0],[3,1,3],[4,1,7],[8,1,7]
19	Nome.x5	1[17,1,8]
20	Nome.9361920)	1[11,1,7]
21	Nome.positivo	3[18,1,3],[4,1,7],[8,1,7]
22	Nome.6,	4[6,1,6],[7,1,6],[4,1,7],[8,1,7]
23	Nome.4g	64[0,1,0],[1,1,0],[2,1,0],[7,1,0],[8,1,0],[10,1,0],[11,1,0],[12,1,0],[14,1,0],[16,1,0],[17,1,0],[10,1,1],[5,1,2],[6,1,2],[9,1,2],[15,1,2],[9,1,3],[15,1,3],[1,1,4],[3,1,4],[0,1,7],[4,1,
24	Nome.wifi	10[0,1,4],[2,1,4],[4,1,7],[8,1,7],[3,1,8],[7,1,8],[13,1,8],[18,1,8],[4,1,9],[11,1,9]
25	Nome.6.0,	23[0,1,1],[2,1,1],[7,1,1],[8,1,1],[12,1,1],[13,1,1],[18,1,1],[0,1,3],[3,1,3],[4,1,3],[6,1,3],[8,1,3],[11,1,3],[12,1,3],[16,1,3],[18,1,3],[19,1,3],[4,1,5],[6,1,5],[10,1,5],[12,1,5],[4,
26	Nome.5.3,	3[2,1,6],[4,1,7],[8,1,7]
27	Nome.6"	5[0,1,0],[2,1,7],[3,1,7],[4,1,7],[8,1,7]
28	Nome.bambu	6[8,1,3],[1,1,4],[3,1,4],[4,1,7],[8,1,7],[9,1,9]

Lista invertida comprimida

FrequênciaTotal[posiçãoCsvGeral, frequênciaLocal]

41,5 kB

1	Preco[1-300]	4[49,1],[72,1],[75,1],[116,1]
2	Preco[301-600]	26[7,1],[9,1],[17,1],[20,1],[24,1],[32,1],[41,1],[43,1],[46,1],[48,1],[68,1],[70,1],[76,1],[82,1],[83,1],[85,1],[97,1],[98,1],[105,1],[108,1],[109,1],[112,1],[113,1],[114,1],[119,1]
3	Preco[601-900]	69[6,1],[10,1],[14,1],[15,1],[18,1],[22,1],[23,1],[26,1],[27,1],[30,1],[31,1],[33,1],[35,1],[36,1],[37,1],[39,1],[40,1],[47,1],[51,1],[52,1],[57,1],[59,1],[64,1],[71,1],[74,1],[77,1]
4	Preco[901-1200]	28[5,1],[8,1],[11,1],[21,1],[25,1],[29,1],[34,1],[44,1],[53,1],[55,1],[66,1],[69,1],[90,1],[93,1],[100,1],[104,1],[106,1],[117,1],[120,1],[121,1],[142,1],[149,1],[150,1],[168,1]
5	Preco[1201-1500]	19[2,1],[19,1],[38,1],[45,1],[61,1],[62,1],[63,1],[79,1],[81,1],[89,1],[91,1],[102,1],[103,1],[143,1],[146,1],[161,1],[167,1],[173,1],[180,1]
6	Preco[1501-1800]	7[0,1],[4,1],[12,1],[16,1],[42,1],[50,1],[163,1]
7	Preco[1801-2100]	
8	Preco[2101-2400]	3[1,1],[13,1],[58,1]
9	Preco[2401-2700]	1[65,1]
10	Preco[2701-3000]	1[60,1]
11	Preco[3001-3300]	
12	Preco[3301-3600]	4[28,1],[54,1],[67,1],[73,1]
13	Preco[3601-3900]	
14	Preco[3901-4200]	1[56,1]
15	Conexao.Wifi	46[0,1],[1,1],[2,1],[3,1],[5,1],[6,1],[7,1],[8,1],[9,1],[11,1],[12,1],[14,1],[15,1],[16,1],[17,1],[125,1],[129,1],[142,1],[143,1],[144,1],[145,1],[146,1],[148,1],[149,1],[150,1],[151,1]
16	Conexao.3g	33[0,1],[1,1],[2,1],[6,1],[7,1],[8,1],[9,1],[12,1],[15,1],[16,1],[17,1],[125,1],[129,1],[142,1],[143,1],[144,1],[145,1],[146,1],[150,1],[153,1],[160,1],[161,1],[166,1],[167,1],[168,1]
17	Conexao.4g	58[0,1],[1,1],[2,1],[3,1],[4,1],[5,1],[6,1],[7,1],[8,1],[10,1],[11,1],[12,1],[13,1],[14,1],[15,1],[16,1],[17,1],[125,1],[129,1],[141,1],[142,1],[143,1],[144,1],[145,1],[146,1],[147,1]
18	Conexao.Bluetooth	2[125,1],[129,1]
19	Conexao.NFC	2[125,1],[129,1]
20	Bateria[1-300]	
21	Bateria[301-600]	
22	Bateria[601-900]	
23	Bateria[901-1200]	
24	Bateria[1201-1500]	
25	Bateria[1501-1800]	
26	Bateria[1801-2100]	1[9,1]
27	Bateria[2101-2400]	4[152,1],[156,1],[157,1],[162,1]
28	Bateria[2401-2700]	11[17,1],[144,1],[147,1],[148,1],[151,1],[154,1],[155,1],[159,1],[163,1],[165,1],[172,1]