

**Universidade de São Paulo
Instituto de Ciências Matemáticas e de Computação
Departamento de Ciências de Computação
Disciplina de Organização de Arquivos
Profa. Dra. Cristina Dutra de Aguiar Ciferri
PAE Fábio Felix Dias (Turma A)
PAE Anderson Chaves Carniel (Turma B)**

**Primeira Parte do Trabalho Prático (Parte I)
Especificação dos Arquivos .csv**

Este documento fornece a descrição sobre os arquivos .csv para sua utilização na Primeira Parte do Trabalho Prático da disciplina de Organização de Arquivos.

Um arquivo .csv é um arquivo texto onde cada linha indica um registro e seus campos são separados pelo carácter “;” (sem aspas). A seguir são detalhados cada arquivo .csv de cada turma. Ao final de cada especificação, é fornecido um exemplo de um registro do respectivo arquivo .csv.

Importante:

- “null” (sem aspas) é usado para indicar que um campo tem valor vazio, ou seja, nulo.
- Os tipos de dados das tabelas não precisam refletir necessariamente o tipo de dados da estrutura em C. Esses tipos são apenas para facilitar o entendimento do arquivo .csv.

Dicionário de Dados – Turma A

(base de dados sobre Domínios de órgãos públicos)

A Tabela 1 fornece as colunas, descrições e os tipos de dados sobre a base de dados dos domínios de órgãos públicos do Brasil. Tal base de dados está contida no arquivo *turmaA-dadosDominios.csv* e a ordem de seus atributos são as mesmas que as mostradas na tabela. É importante notar que o arquivo .csv não contém cabeçalho, então é estritamente necessário seguir a ordem usada na tabela.

Tabela 1. Caracterização do arquivo turmaA-dadosDominios.csv

Nome do Campo	Ordem do Campo no Arquivo .csv	Tipo de Dado no .csv	Descrição
<i>dominio</i>	1	Texto (tamanho variável)	URL cadastrada no Registro.br
<i>documento</i>	2	Texto (tamanho fixo)	Número do documento informado ao Registro.br no cadastro da URL. Este campo tem a seguinte máscara: #####."#####"/#####-#### (e.g. 000.394.411/0001-09)
<i>nome</i>	3	Texto (tamanho variável)	Nome do Órgão ou Entidade responsável pela URL
<i>cidade</i>	4	Texto (tamanho variável)	Cidade cadastrada no Registro.br
<i>uf</i>	5	Texto (tamanho variável)	Nome do estado por extenso
<i>dataHoraCadastro</i>	6	Texto/Timestamp (tamanho fixo)	Data e hora de quando a URL foi cadastrada. Este campo tem a seguinte máscara: dd/mm/aaaa hh:mm:ss (e.g., 17/12/2004 17:03:07)
<i>dataHoraAtualiza</i>	7	Texto/Timestamp (tamanho fixo)	Data e hora de quando a URL foi atualizada. Este campo tem a seguinte máscara: dd/mm/aaaa hh:mm:ss (e.g., 17/12/2004 07:03:07)
<i>ticket</i>	8	Numérico (tamanho fixo)	Número do ticket vinculado ao registro da URL no cadastro do Registro.br

Exemplo de registro:

A primeira linha do arquivo .csv segue abaixo:

eletrobras.gov.br;000.001.180/0002-07;CENTRAIS ELETRICAS BRASILEIRAS SA;Rio de Janeiro;Rio de Janeiro;09/01/1996 12:00:00;02/07/2014 19:56:36;5023

O valor “eletrobras.gov.br” refere-se ao campo “*dominio*”, o valor “000.001.180/0002-07” refere-se ao campo “*documento*”, e assim sucessivamente para os outros campos, conforme a ordem mostrada na tabela 1.

Dicionário de Dados – Turma B

(base de dados sobre Companhias do Brasil)

A Tabela 2 fornece as colunas, descrições e os tipos de dados sobre a base de dados das companhias do Brasil. Tal base de dados está contida no arquivo *turmaB-dadosCompanhias.csv* e a ordem de seus atributos são as mesmas que as mostradas na tabela. É importante notar que o arquivo .csv não contém cabeçalho, então é estritamente necessário seguir a ordem usada na tabela.

Tabela 2. Caracterização do arquivo turmaB-dadosCompanhias.csv

Nome do Campo	Ordem do Campo no Arquivo .csv	Tipo de Dado no .csv	Descrição
<i>CNPJ</i>	1	Texto/Numérico (tamanho fixo)	CNPJ de cadastro da companhia. Este campo possui a seguinte máscara: ##“.”####“.”####“/”####“-”## (e.g. 44.597.052/0001-62).
<i>nomeSocial</i>	2	Texto (tamanho variável)	Nome no qual foi registrado o CNPJ.
<i>nomeFantasia</i>	3	Texto (tamanho variável)	Nome fantasia da companhia.
<i>dataRegistro</i>	4	Texto/Data (tamanho fixo)	Data em que a companhia foi registrada sob o CNPJ. Este campo possui a seguinte máscara: dd/mm/aaaa.
<i>dataCancelamento</i>	5	Texto/Data (tamanho fixo)	Data em que a companhia foi fechada. Este campo possui a seguinte máscara: dd/mm/aaaa.
<i>motivoCancelamento</i>	6	Texto (tamanho variável)	Razão pela qual a companhia foi fechada.
<i>nomeEmpresa</i>	7	Texto (tamanho variável)	Nome da empresa que presta auditoria.
<i>CNPJauditor</i>	8	Texto/Numérico (tamanho fixo)	CNPJ da empresa que presta auditoria para a companhia. Este campo possui a seguinte máscara: ##“.”####“.”####“/”####“-”## (e.g., 44.597.052/0001-62).

Exemplo de registro:

A primeira linha do arquivo .csv segue abaixo:

11.396.633/0001-87;3A COMPANHIA SECURITIZADORA;TRIPLO A COMPANHIA SECURITIZADORA;08/03/10;18/12/15;Cancelamento Voluntário - In Cvm 480/09;MOORE STEPHENS LIMA LUCCHESI AUDITORES INDEPENDENTES;60.525.706/0001-07

O valor “11.396.633/0001-87” refere-se ao campo “*CNPJ*”, o valor “3A COMPANHIA SECURITIZADORA” refere-se ao campo “*nomeSocial*”, e assim sucessivamente para os outros campos, conforme a ordem mostrada na Tabela 2.