Atividade de Laboratório II

(Prazo para conclusão da atividade: 23/03/2017)

1. Faça um *fork* (por meio do github) do projeto *pythoncourse*. Em seguida, faça um *clone* deste projeto em sua máquina. Abra o projeto utilizando a IDE LiClipse (ou Eclipse). No *fork* do projeto original, crie um *branch* chamado *bugfix\_production*.
2. Melhore o método *extract\_filename*, no arquivo *utils*. O método original supõe que o nome do arquivo terá apenas um ponto(.). Mas sabemos que um usuário pode nomear arquivos com vários pontos. Por exemplo: “meu.dados.txt”. Utilizando a implementação atual de *extract\_filename* na *string* “meu.dados.txt”, o resultado é simplesmente a *string* “meus”, quando deveria ser *“meus.dados”*. Corrija este problema e faça um *commit*.
3. Depois de fazer todas as correções sugeridas e subir os dados para o servidor (com *push*) para o *branch* *bugfix\_production*, faça um *merge* deste *branch* com o *branch* *master*. Em seguida, faça um *pull request* para o repositório original.
4. Crie um novo *branch* chamado indexao\_dados. Neste *branch*, implemente uma nova funcionalidade. A funcionalidade consiste no carregamento de informações sobre unidades de saúde, cujos dados estão disponíveis no seguinte link: <http://repositorio.dados.gov.br/saude/unidades-saude/unidade-basica-saude/ubs.csv.zip>. Primeiramente, implemente uma função que carrega os dados em uma lista, onde cada elemento da lista contém uma *tupla*. As linhas do arguivo *csv* baixados (dentro do ubs.csv.zip) contém valores separados por vírgula. A primeira *tupla* na lista deve conter todos os valores da primeira linha; o segundo elemento, todos os valores da segunda linha; e assim por diante. O nome o método é *loadlistfromcsv*, que recebe a URL do arquivo e retorna uma lista com o formato descrito. Realize um *commit* seguido de um *push* para sincronizar o repositório local com o repositório remoto.
5. No *branch indexao\_dados*, crie um método (chamado de *create\_cidcnes\_index*) que recebe a lista produzida por *loadlistfromcsv* e, com base nessa lista, retorna um dicionário que tem como chave o código do município concatenado com código CNES da unidade de saúde. O valor de cada elemento do dicionário é uma par *chave:valor*, onde chave é a concatenação do código da cidade com o código CNESe *valor* é a *tupla* da lista retornada pelo método *loadlistfromcsv* que tenha o código da cidade e o código CNEScorrespondente ao código da chave.

Exemplo, suponha que o arquivo tenha os seguintes dados:

vlr\_latitude,vlr\_longitude,cod\_munic,cod\_cnes,nom\_estab,dsc\_endereco,dsc\_bairro,dsc\_cidade,dsc\_telefone,dsc\_estrut\_fisic\_ambiencia,dsc\_adap\_defic\_fisic\_idosos,dsc\_equipamentos,dsc\_medicamentos

-10.9112370014188,-37.0620775222768,280030,3492,US OSWALDO DE SOUZA,TV ADALTO BOTELHO,GETULIO VARGAS,Aracaju,7931791326,Desempenho acima da média,Desempenho muito acima da média,Desempenho mediano ou um pouco abaixo da média,Desempenho acima da média

-9.48594331741306,-35.8575725555409,270770,6685315,USF ENFERMEIRO PEDRO JACINTO AREA 09,R 15 DE AGOSTO,CENTRO,Rio Largo,Não se aplica,Desempenho mediano ou um pouco abaixo da média,Desempenho mediano ou um pouco abaixo da média,Desempenho mediano ou um pouco abaixo da média,Desempenho mediano ou um pouco abaixo da média

Este arquivo tem 3 linhas.

Linha 1: vlr\_latitude,vlr\_longitude,cod\_munic,cod\_cnes,nom\_estab,dsc\_endereco,dsc\_bairro,dsc\_cidade,dsc\_telefone,dsc\_estrut\_fisic\_ambiencia,dsc\_adap\_defic\_fisic\_idosos,dsc\_equipamentos,dsc\_medicamentos

Linha 2:

-10.9112370014188,-37.0620775222768,280030,3492,US OSWALDO DE SOUZA,TV ADALTO BOTELHO,GETULIO VARGAS,Aracaju,7931791326,Desempenho acima da média,Desempenho muito acima da média,Desempenho mediano ou um pouco abaixo da média,Desempenho acima da média

Linha 3:

-9.48594331741306,-35.8575725555409,270770,6685315,USF ENFERMEIRO PEDRO JACINTO AREA 09,R 15 DE AGOSTO,CENTRO,Rio Largo,Não se aplica,Desempenho mediano ou um pouco abaixo da média,Desempenho mediano ou um pouco abaixo da média,Desempenho mediano ou um pouco abaixo da média,Desempenho mediano ou um pouco abaixo da média

Observe que cada linha tem treze valores separados por vírgula. A primeira linha tem o código dos campos que descrevem os valores nas linhas seguintes. Dessa forma, o método *loadlistfromcsv* deve receber o *caminho* (ou *path*) do arquivo, lê o conteúdo do arquivo como *texto* e produzir uma lista de tuplas, tal que cada tupla tenha treze elementos correspondentes aos valores dos campos. A função *loadlistfromcsv* aplicada a um arquivo com o conteúdo acima, deveria retornar a seguinte lista:

[(‘vlr\_latitude’,’vlr\_longitude’,’cod\_munic’,’cod\_cnes’,’nom\_estab’,’dsc\_endereco’,’dsc\_bairro’,’dsc\_cidade’,’dsc\_telefone’,’dsc\_estrut\_fisic\_ambiencia’,’dsc\_adap\_defic\_fisic\_idosos’,’dsc\_equipamentos’,’dsc\_medicamentos’),( '-10.9112370014188’,’-37.0620775222768’,’280030’,’3492’,’US OSWALDO DE SOUZA’,’TV ADALTO BOTELHO’,’GETULIO VARGAS’,’Aracaju’,’7931791326’,’Desempenho acima da média’, ’Desempenho muito acima da média’,’Desempenho mediano ou um pouco abaixo da média’,’Desempenho acima da média’), (‘-9.48594331741306’,’-35.8575725555409’,’270770’,’6685315’,’USF ENFERMEIRO PEDRO JACINTO AREA 09’,’R 15 DE AGOSTO’,’CENTRO’,’Rio Largo’,’Não se aplica’,’Desempenho mediano ou um pouco abaixo da média’,’Desempenho mediano ou um pouco abaixo da média’,’Desempenho mediano ou um pouco abaixo da média’,’Desempenho mediano ou um pouco abaixo da média’)]

Observe que todos os valores das tuplas são lidos como *strings*. Já o método create\_cidnes\_index aplicado a esta lista, resultaria no seguinte dicionário:

{ **’2800303492’**:( '-10.9112370014188’,’-37.0620775222768’,’280030’,’3492’,’US OSWALDO DE SOUZA’,’TV ADALTO BOTELHO’,’GETULIO VARGAS’,’Aracaju’,’7931791326’,’Desempenho acima da média’, ’Desempenho muito acima da média’,’Desempenho mediano ou um pouco abaixo da média’,’Desempenho acima da média’), **‘2707706685315’**: (‘-9.48594331741306’,’-35.8575725555409’,’270770’,’6685315’,’USF ENFERMEIRO PEDRO JACINTO AREA 09’,’R 15 DE AGOSTO’,’CENTRO’,’Rio Largo’,’Não se aplica’,’Desempenho mediano ou um pouco abaixo da média’,’Desempenho mediano ou um pouco abaixo da média’,’Desempenho mediano ou um pouco abaixo da média’,’Desempenho mediano ou um pouco abaixo da média’)}