

NOM:

Prénom:

Examen de « Visualisation Interactive de Données »

Enseignant : Romain Vuillemot

Aucun document (support de cours, TD, etc.) n'est admis (sauf dictionnaires de langues pour les étrangers).

Les ordinateurs sont interdits sauf les traducteurs électroniques.

-En cas de question / interrogation sur le sujet, annotez le (en rajoutant du texte) afin d'indiquer vos hypothèses ou interprétations (dans la limite du raisonnable).

-Il est préférable de faire des réponses courtes et synthétiques, plutôt que longues et verbeuses.

-Ecrire lisiblement, éviter les ratures et si vous introduisez des acronymes (ex : DOM, JS, ..) définissez-les avant.

-Pour la partie conception graphique, utilisez ou non des couleurs, mais si vous n'en n'avez pas ce n'est pas un soucis (utilisez des hachures ou autres propriétés visuelles pour simuler les couleurs si besoin)

A - Questions générales (4 points)

(2pt) Qu'est-ce que l' « interaction » dans le contexte d'une visualisation interactive de données ?

Indiquez pourquoi cela est parfois nécessaire et donnez 3 exemples d'interactions avec une visualisation.

(1pt) Détaillez un exemple d'interaction efficace pour une visualisation en particulier (que vous pouvez illustrer)

(1pt) A l'inverse, détaillez un exemple d'interaction inéficace et indiquez quelle serait l'alternative.

B – Questions techniques sur Observable/D3.JS (6 points)

(2pt) Que permet de réaliser la fonction **d3.group** en Observable/D3.JS ? Comment peut-on encoder le résultat de cette fonction dans une visualisation ?

(2pt) Que fait le code Observable/D3.JS de la page suivante ? Dessinez le résultat attendu (tel qu'affiché par le navigateur), et donnez la structure du DOM (il n'est pas nécessaire de la détailler, mais indiquer l'arborescence principale) qui correspond au SVG produit..

(1pt) Indiquez sur le code de la page suivante :

- Où est réalisé le data-binding ?
- Ce qu'affiche la commande `console.log(data)` ?

(1pt) Les données utilisées dans la page suivantes sont aléatoires. Proposer un échantillon de jeu de données qui pourrait remplacer ces données aléatoires et ainsi être visualisé d'une façon quasi identique (proposez ce jeu de données sous forme de table comme pour la Question D, avec : des types de données, quelques valeurs, mais pas un jeu de données complet).

```

chart = {

  var margin = {top: 20, right: 10, bottom: 20, left: 10};

  const width = 600 - margin.left - margin.right,
    height = 100 - margin.top - margin.bottom;

  const svg = d3.create("svg")
    .attr("viewBox", [0, 0, width, height]);

  var n = 5, m = 3;

  var data = d3.range(m).map(function() {
    return d3.range(n).map(Math.random);
  });

  console.log(data);

  const c = d3.scaleOrdinal(d3.schemeCategory10).domain(d3.range(n));

  const x = d3.scaleBand().domain(d3.range(m)).range([0, width]).paddingInner(.1);

  const x2 = d3.scaleBand().domain(d3.range(n)).range([0, x.bandwidth()]).paddingInner(.1);

  var y = d3.scaleLinear().domain([0, 1]).range([0, height]);

  svg.selectAll("g").data(data).enter()
    .append("g")
    .attr("transform", function(d, i) { return "translate(" + x(i) + ", 0)"; })
    .selectAll("rect")
    .data(function(d) { return d; }).enter()
    .append("rect")
    .style("stroke", "black")
    .attr("fill", function(d, i) { return c(i); })
    .attr("x", function(d, i) { return x2(i); })
    .attr("y", function(d, i) { return height - y(d); })
    .attr("width", function(d, i) { return x2.bandwidth(); })
    .attr("height", function(d, i) { return y(d); })

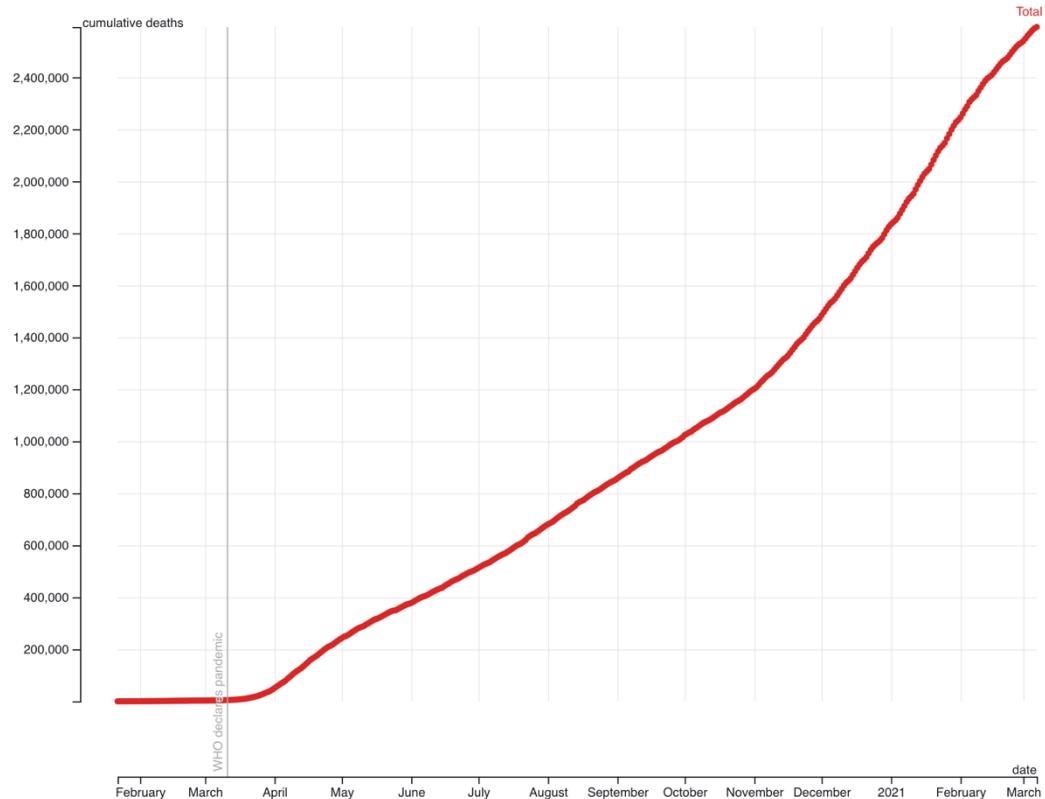
  return svg.node();
}

```

C – Cas d'étude : critique de design et re-design (5 points)

Ci-dessous un graphique temporel lié à la propagation de la COVID-19 où la courbe indique le nombre de déces cumulés de février 2020 à début mars 2021 dans le monde entier.

NB: le graphique a une échelle verticale linéaire.



(1pt) Quels sont les avantages et inconvénients de cette visualisation ?

(2pt) Proposer de nouvelles visualisation permettant de combler un ou plusieurs inconvénients que vous avez identifié et motivez votre choix.

(2pt) En particulier indiquez les traitements de données nécessaires pour vos propositions de visualisations (ex: filtrage, aggregation, enrichissement de jeu de données).

D – Cas d'étude : design d'une visualisation (5 points)

Nous vous fournissons un extrait de jeu de données lié aux passagers du Titanic (un bateau qui a coulé le 15 avril 1912 avec 1310 personnes à bord).

pclass	survived	name	sex	age	sibsp	parch	ticket	fare	cabin	embarked	home.dest
1	1	Allen, Miss. Elisabeth W	female	29	0	0	24160	211.3375	B5	S	St Louis, MO
1	1	Allison, Master. Hudson	male	0.9167	1	2	113781	151.55	C22 C26	S	Montreal, PQ / Chesterville, ON
1	0	Allison, Miss. Helen Lora	female	2	1	2	113781	151.55	C22 C26	S	Montreal, PQ / Chesterville, ON
1	0	Allison, Mr. Hudson Josh	male	30	1	2	113781	151.55	C22 C26	S	Montreal, PQ / Chesterville, ON
1	0	Allison, Mrs. Hudson J C	female	25	1	2	113781	151.55	C22 C26	S	Montreal, PQ / Chesterville, ON
1	1	Anderson, Mr. Harry	male	48	0	0	19952	26.55	E12	S	New York, NY
1	1	Andrews, Miss. Kornelia	female	63	1	0	13502	77.9583	D7	S	Hudson, NY
1	0	Andrews, Mr. Thomas Jr	male	39	0	0	112050		0 A36	S	Belfast, NI
2	0	Wheaton, Mr. Edward H	male	66	0	0	C.A. 24579	10.5		S	Guernsey, England / Edgewood, RI
2	0	Wheeler, Mr. Edwin "Freddie"	male	29.8811	0	0	SC/PARIS 21	12.875		S	
2	1	Wilhelms, Mr. Charles	male	31	0	0	244270		13	S	London, England
2	1	Williams, Mr. Charles E	male	29.8811	0	0	244373		13	S	Harrow, England
2	1	Wright, Miss. Marion	female	26	0	0	220844	13.5		S	Yoevil, England / Cottage Grove, OR
2	0	Yrois, Miss. Henriette ("M	female	24	0	0	248747		13	S	Paris
3	0	Abbing, Mr. Anthony	male	42	0	0	C.A. 5547	7.55		S	
3	0	Abbott, Master. Eugene	male	13	0	2	C.A. 2673	20.25		S	East Providence, RI
3	0	Abbott, Mr. Rossmore E	male	16	1	1	C.A. 2673	20.25		S	East Providence, RI
3	1	Abbott, Mrs. Stanton (Ro	female	35	1	1	C.A. 2673	20.25		S	East Providence, RI
3	1	Abelseth, Miss. Karen M	female	16	0	0	348125	7.65		S	Norway Los Angeles, CA
3	1	Abelseth, Mr. Olaus Jorg	male	25	0	0	348122	7.65	F G63	S	Perkins County, SD

Voici la description des variables (colonnes du tableau du jeu de données) :

pclass : 1eme, 2eme ou 3eme classe survived : Oui 1, Non 0 name : Nom et prénom sex : Sexe age : Age sibsp : Nb de frères/soeurs ou époux	parch : Nombre de parents / enfants ticket : Numéro du ticket fare : Prix du ticket (en Livre Sterling) cabin : Numéro de cabine embarked : Port d'embarquement C = Cherbourg; Q = Queenstown; S = Southampton home.dest : Lieu de destination
--	---

Votre rôle est de répondre aux questions ci-dessous au moyen de visualisations de données. A noter que du fait que le jeu de données soit partiel, vous êtes libres de faire des hypothèses sur la distribution des données (et parfois même du type si besoin).

Questions :

(1pt) Pour chaque variable (colonne) indiquez le type de données (directement sur la table ci-dessus)

(1pt) Dessinez une visualisation montrant la répartition des passagers par taux de survivant.

(1pt) Comment visualiser plusieurs dimensions comme classe, sexe, et si survécu en une même visualisation ? Proposez un exemple de visualisation.

(2pt) Quelle visualisation supplémentaire proposeriez-vous ? Motivez et illustrez deux visualisations et quelles sont les questions auxquelles elles répondent. (si vous avez besoin de données ou traitement de données supplémentaires, indiquez quelles colonnes sont impliquées)