**Meilleures réponses aux questions de l'examen professionnel
pour l'accès au corps d'attaché statisticien de l'Insee**

Année 2024

Sujet N°1 : Domaine démographique et social

Vous trouverez ci-après une sélection faite par le jury à l'issue de la correction de l'épreuve écrite. Elle vise notamment à éclairer les candidats sur les attendus en termes de connaissances et de longueur de la réponse.

Questions de cadrage (6 points)

1 -Au sein du SSP, certaines enquêtes sont réalisées en panel. Donnez un intérêt et un inconvénient de ce protocole d'enquête. Citez deux enquêtes en panel.

Une enquête panel est une enquête qui est réalisée plusieurs fois dans le temps auprès des mêmes personnes. L'intérêt de ce type d'enquête est que ça permet de suivre au fil du temps l'évolution des individus par rapport au sujet abordé. L'un des inconvénients de ce protocole d'enquête est que les questionnements successifs sur les mêmes personnes entraînent une charge pour l'enquêté (des fois sur plusieurs années) et donc cela entraîne souvent un effet d'attrition, c'est-à-dire qu'il y a de moins en moins de répondants au fil des questionnements. Il n'est pas toujours simple de motiver les personnes à répondre plusieurs fois.

Il existe plusieurs enquêtes panel dans le SSP : on peut citer l'enquête Entrée dans la vie Adulte ou l'enquête Patrimoine.

2 -Dans le graphique 1, l'auteure présente des intervalles de confiance. Expliquez ce qu'ils représentent et comment les utiliser pour comparer deux résultats.

Les intervalles de confiance permettent de représenter le biais qu'engendre une enquête par échantillon par rapport à l'ensemble de la population. Cela représente la marge d'erreur. Pour un intervalle de confiance à 95 %, la réalité de la statistique calculée est à 95 % de chances de se trouver dans l'intervalle.

Pour comparer deux résultats avec des intervalles de confiance, il faut vérifier si le minimum et le maximum de chacun des résultats ne se chevauchent pas. Si c'est le cas, la comparaison ne peut pas se faire. La comparaison n'est pas significative.

3 -A quoi sert une régression linéaire dans une étude statistique (cinq lignes environ) ? Donnez une limite dans l'utilisation de cette méthode.

Une régression linéaire permet de définir un lien mathématique entre plusieurs observations d'une variable. Elle permet ainsi de définir une fonction représentative pour définir une variable en fonction d'une autre. Cette méthode peut avoir comme limite l'existence de valeurs extrêmes ou aberrantes qui vont fausser l'estimation des paramètres de la fonction recherchée.

4 - Donnez deux bonnes pratiques dans la conception d'une question pour une enquête. Pour chacune d'elles, quel est l'impact de son non-respect ?

Il faut tout d'abord utiliser des mots simples et facilement compréhensibles, y compris pour des personnes maîtrisant peu la langue française. Ne pas respecter cette règle augmente le risque de non-réponse à cette enquête ou des réponses erronées.

Il faut également poser de préférence des questions fermées avec des listes de modalités courtes. Ne pas respecter cette règle peut compliquer l'exploitation de l'enquête. Une liste de modalités trop longues peut également fausser les résultats en sur-représentant les premières ou dernières modalités.

5 - Des enquêtes ménages ont été réalisées pendant le confinement. Selon vous, quel est le mode de collecte le plus adapté dans ces circonstances ? Expliquez pourquoi il est pertinent et donnez une limite de ce mode de collecte.

Bonne réponse possible pour la réponse par internet

Le mode de collecte le plus adapté en cas de confinement est la collecte par internet. Ce mode de collecte est pertinent car les enquêteurs n'ont pas besoin de se déplacer pour voir les enquêtés. Les enquêtés peuvent également répondre tranquillement sur leur ordinateur à l'heure de leur choix en dérangeant le moins possible les autres personnes confinées du ménage. En revanche, le mode de collecte par internet suppose d'être à l'aise avec les outils numériques et de bien comprendre et répondre aux questions. Cela exclut une partie de la population qui n'a pas internet ou n'est pas à l'aise ou les personnes illettrées.

Autre bonne réponse possible pour la réponse par téléphone

Le mode le plus adapté est le téléphone. Ce mode de collecte permet de limiter le contact physique entre enquêteurs et enquêtés, mais permet également de conserver l'accompagnement des enquêteurs lors de la réponse à l'enquête (explication de concept, normalisation des réponses ouvertes comme la profession...). Il est également plus facile à mettre en place qu'un questionnaire sur internet.

Par contre, ce mode de collecte augmente les difficultés de prise de contacts des enquêteurs. En effet, il faut récupérer un numéro de téléphone qui n'est pas toujours disponible ou à jour dans les bases de sondage ? Les refus sont également plus nombreux et faciles par téléphone car les enquêtés confondent les enquêtes de l'Insee avec du démarchage téléphonique.

Questions de compréhension et de connaissance du sujet (7 points)

6 - Le tableau 1 montre le résultat de la régression linéaire sur le score de dépression sur l'ensemble des personnes en emploi. Pouvez-vous expliquer simplement le coefficient « 4,0 » pour la modalité « Femme » ?

Toutes choses égales par ailleurs, être une femme a un fort effet sur le score de dépression : + 4 points comparé aux hommes.

7 -En vous appuyant sur le texte, donnez les valeurs des cases A, B, C, D et E grisées dans le tableau 1.

A = 3,0 ; B = 46 % ; C = 54 % ; D = 72 % ; E = 28 %

8 -Dans le graphique 2, proposez une note de lecture qui explique les résultats observés pour les personnes déclarant se sentir fatiguées ou manquer d'énergie.

Une note de lecture pourrait être : 75 % des personnes salariées de l'hôpital déclarent avoir été gênées au moins à plusieurs reprises lors des quinze derniers jours par de la fatigue ou un manque d'énergie, contre 61 % pour l'ensemble des personnes en emploi.

9 -En vous appuyant sur le graphique 3, calculez la contribution du sexe à la différence de score moyen de dépression entre le personnel hospitalier et l'ensemble des personnes en emploi. Vous explicitez la méthode de calcul.

La différence de scores de dépression entre personnels hospitaliers et ensemble des personnes en emploi est égale à la somme des contributions, c'est-à-dire à : Contribution sexe + Contribution autres caractéristiques personnelles + Contribution travail à l'hôpital + Contribution conditions de travail difficiles

Soit :

$$+ 2,4 = x + 0,0 - 1,3 + 2,5$$

$$- x = 2,5 - 1,3 - 2,4$$

$$- x = 0,1 - 1,3 = - 1,2$$

$$X = + 1,2$$

La contribution du sexe à la différence de score de dépression entre personnel hospitalier et ensemble des personnes en emploi est donc égale à + 1,2 points.

10 -L'auteure montre qu'une fois prises en compte les conditions de travail et les caractéristiques personnelles, le fait de travailler à l'hôpital diminue le risque de développer des symptômes de la dépression. Donnez deux hypothèses formulées par l'auteure expliquant ce résultat.

Travailler à l'hôpital durant la crise sanitaire diminue le score moyen de dépression (- 1,3). Une raison à cela est la présence de ressources à l'hôpital permettant de limiter les risques psychosociaux. L'entraide entre collègues vivant les mêmes situations difficiles, le sentiment d'être utile en participant à sauver des vies, concourent à limiter les risques psychosociaux. Un autre facteur d'une exposition plus modérée à la dépression pour le personnel hospitalier est leur résilience qui est supérieure à l'ensemble de la population, de par la nature même de leur travail.

11 -Rédiger, en 10 à 15 lignes, un résumé à destination du grand public.

Le troisième volet de l'enquête EpiCov s'est déroulé en 2021 et a permis de présenter des indicateurs liés à la dépression et à l'anxiété, en comparant le personnel hospitalier à l'ensemble des personnes en emploi. Le score moyen de dépression des personnels hospitaliers est plus élevé que pour l'ensemble : + 2,4 points. Le rôle des conditions de travail sur la santé est démontré, notamment par les risques psychosociaux qu'ils engendrent. De façon générale, les facteurs les plus impactants sont : d'être une femme, avoir des difficultés à concilier la vie personnelle et profession-

nelle, et subir une pression pour ne pas s'arrêter. Le fait que le personnel hospitalier est plus touché par les symptômes de dépression peut s'expliquer par des conditions de travail difficiles, et par la forte féminisation (78 %) du personnel hospitalier.

Questions de connaissance de la statistique publique (7 points)

- 12 -Les dispositifs de l'enquête emploi en continu (EEC), du recensement de la population (RP) et des demandeurs d'emploi en fin de mois (DEFM) permettent d'éclairer la thématique du chômage. Expliquez les différences de concepts et de mesures entre ces sources en 10 à 15 lignes ?

L'enquête emploi en continu est la seule enquête permettant de répondre au critère de chômage au sens du bureau international du travail (BIT) permettant des comparaisons entre les pays. Pour être considérée comme une personne au chômage les critères sont strictes selon une semaine de référence donnée et cette enquête permet notamment de quantifier le halo du chômage et le sous-emploi mais du fait de la taille de son échantillon les données ne permettent pas d'être exploitées à un niveau trop fin comme la maille communale. Contrairement au recensement qui permet des analyses plus fines mais non conjoncturelles et qui reposent uniquement sur un questionnaire auto administré. Il s'agit de données déclaratives et uniquement annuelles avec un certain laps de temps écoulé. En ce qui concerne les demandeurs d'emploi en fin de mois, il s'agit d'une source administrative qui ne prend que les personnes inscrites à France travail (ex-Pole Emploi) qui sont ensuite classées selon certaines catégories. Cependant il est possible d'être chômeur au sens du BIT sans être demandeur d'emploi et inversement. Les DEFM sont aussi mobilisables à une échelle fine et pour chacun des trimestres.

- 13 -L'Insee et le SSP se dotent progressivement d'une nouvelle filière d'enquête. Après avoir rappelé le champ des enquêtes concernées, donnez et illustrez concrètement trois objectifs de ce programme.

La nouvelle filière de collecte d'enquête va accueillir à la fois les enquêtes entreprises, ménages et elle sera également ouverte aux SSM. Voici 3 objectifs pour ce programme :

- un pilotage par la métadonnée active qui s'appuie sur le répertoire de métadonnées statistiques (Rmès), qui permet de s'appuyer sur des concepts harmonisés et communs à l'ensemble de la statistique ainsi que le recours automatique aux nomenclatures (PCSn APE, ...)
- un travail en mode projet avec des moyens dédiés et rationalisés dans le domaine métier et informatique de la conception à la production puis la maintenance ;
- une forte mutualisation des connaissances des différents domaines sur la question du multimode, notamment en découpant les questionnaires en « cubes », en modules, facilement réutilisables d'une enquête à l'autre.

- 14 -Dans le cadre de la multiplication des cyberattaques, indiquez une mesure prise par les équipes informatiques pour protéger le système d'information et son intérêt. Quelles sont les bonnes pratiques à adopter au quotidien par les agents du SSP pour se prémunir d'une cyberattaque (donnez 3 exemples) ?

Afin de limiter le risque, pour le système d'information de l'INSEE, liées aux cyberattaques, les équipes informatiques utilisent un système pointu d'antivirus, d'antimalwares et autres dispositifs techniques qui permettent le repérage et la neutralisation de ces attaques via l'informatique et le réseau internet. La mise en place du plan de reprise d'activité, récemment mis en place permet, quoi qu'il arrive, de pouvoir réaliser nos travaux (c'est une « copie » du site de production d'Osny installé à Toulouse - Auzeville qui prendrait le relai en cas de besoin). Au quotidien, les équipes in-

formatiques nationales et locales rappellent aux agents du SSP certaines bonnes pratiques à respecter pour limiter le risque : ne pas divulguer un mot de passe, verrouiller son poste de travail si on le quitte, ne pas ouvrir un mail dont on ne connaît l'expéditeur et encore moins cliquer sur un lien présent dans ce mail.

15 -Le plan « Services publics écoresponsables » (SPE) a été lancé en 2020. Citez trois objectifs de développement durable pour l'administration. Si vous étiez responsable SPE de votre service, quelle serait la première action que vous proposeriez et comment procéderiez-vous pour la mettre en place ?

Le SPE lancé en 2020 est appliqué au niveau du ministère de l'Economie et des finances via le projet « Bercy Vert » puis à l'Insee via le projet « Insee Vert ».

L'objectif principal est de limiter, voire réduire l'empreinte carbone des administrations publiques. On peut citer 3 objectifs :

- repenser les déplacements des agents. On peut citer la mise en place du forfait mobilité durables pour les agents qui viennent à vélo ou utilisent le covoiturage. Il permet de réduire les GES.
- repenser la restauration collective des agents : proposer de la nourriture biologique, locale, des repas végétariens.
- diminuer la consommation d'énergie des établissements : baisse du chauffage, éteindre les lumières, etc.

Si j'étais responsable SPE d'un service, ma première action serait de réduire mon empreinte carbone de nos outils de travail notamment nos données numériques (archivage des données, suppression, ménage des BAL fonctionnelles voire individuelles). Pour ce faire, j'organiserais des points d'information, de sensibilisation et j'impliquerais l'équipe à la réflexion des bonnes pratiques déclinables en fiches pratiques ou charte.